

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 1 043 399 A2**

(12)

**EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:  
11.10.2000 Patentblatt 2000/41

(21) Anmeldenummer: 00105929.4

(22) Anmeldetag: 23.03.2000

(51) Int. Cl.<sup>7</sup>: **C12N 15/86**, C12N 7/01,  
C12N 7/04, C12N 5/10,  
C07K 14/18, A61K 49/00,  
A61K 48/00

(84) Benannte Vertragsstaaten:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU**  
**MC NL PT SE**  
Benannte Erstreckungsstaaten:  
**AL LT LV MK RO SI**

(30) Priorität: 03.04.1999 DE 19915178

(71) Anmelder:  
**Bartenschlager, Ralf, Dr.**  
**55239 Gau-Odernheim (DE)**

(72) Erfinder:  
**Bartenschlager, Ralf, Dr.**  
**55239 Gau-Odernheim (DE)**

(74) Vertreter: **Rudolph, Ulrike, Dr.**  
**Patentanwältin**  
**In der Schanz 10**  
**69198 Schriesheim (DE)**

(54) **Hepatitis C Virus Zellkultursystem**

(57) Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepato-mazellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

**EP 1 043 399 A2**

## Beschreibung

[0001] Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiziert sind.

[0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

[0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989: Science, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartschläger 1997, Intervirology 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozessierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

[0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3' NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

[0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschrittsmäßig bei allen Blutkonserven angewendet werden (müssen).

[0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Kompletsequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

[0007] Die klonierten und charakterisierten Teil- und Kompletsequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, Cell, 87,343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463; Cho et al., 1998, J. Biol. Chem., 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

[0008] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

[0009] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, Proceedings of the National Academy of Sciences, USA, 90, 6037-6041; Mizutani et al. 1996, Journal of Virology, 70, 7219-7223; M. Ikeda et al. 1998, Virus Res. 56, 157; Fournier et al. 1998, J. Gen. Virol. 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, Journal of General Virology, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Proteine in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

[0010] Desweiteren ist aus den Publikationen von Yoo et al. (1995, Journal of Virology, 69, 32-38) und von Dash et al., (1997, American Journal of Pathology, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels in vitro Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeführten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.

[0011] Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines HCV- Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.

[0012] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das im

wesentlichen die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

**[0013]** Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

**[0014]** Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

**[0015]** Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markergen (Selektionsgen) umfaßt.

**[0016]** Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

**[0017]** Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenese und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

**[0018]** Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

**[0019]** Eine Ausführungsform des HCV-RNA-Konstrukts, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO:1 umfaßt.

Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO:2 oder SEQ ID NO:3 oder SEQ ID NO:4 oder SEQ ID NO:5 oder SEQ ID NO:6 oder SEQ ID NO:7 oder SEQ ID NO:8 oder SEQ ID NO:9 oder SEQ ID NO:10 oder SEQ ID NO:11 umfassen.

**[0020]** Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3' NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

- 5 (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA  
GCTTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
AGACTGGAGT
- 10 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCCAGACT GGAGC
- 15 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCCAGACT GCAGC
- 20 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCAG ACTGGAGT
- 25 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCAG ACTGGAGT
- 30 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTT TTGAGAGAGA GAGTCTCACT  
CTGTTGCCA GACTGGAGT
- 35 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTT GAGAGAGAGA  
GTCTCACTCT GTTGCCCAGA CTGGAGT
- 40 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTTAAT CTTTTTTTCT TTTTCCTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 45 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTTAATC TTTTTTTTCT TTTCTTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

50 Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blasticidin.

55 [0021] Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

[0022] Eine Alternative zu den Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-

Selektion durchgeführt werden kann.

[0023] Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5' NTR, d.h. strangabwärts der 5' NTR bzw. strangaufwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion

im Bereich der 3' NTR oder an anderer Stelle des HCV-Genoms oder -Subgenoms, z.B. innerhalb des Polyproteins.  
[0024] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

[0025] Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0026] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

[0027] Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

[0028] Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

[0029] Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

[0030] Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

[0031] Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms einkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polyproteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'high-throughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.

- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflusst werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- 5 • Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder  
10 anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausch/s in das Ursprungs-konstrukt kann deren  
15 Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von  
20 Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellkultursystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb in der Lage infektiöse Viren zu produzieren.
- 25 **[0033]** Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:
  - Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen: Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertionen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringerter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
  - Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als  
30 leberzellspezifische Genföhrn in der Gentherapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
  - Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das allerdings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumsselektion unterzogen oder,  
45 in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine Antibiotikumsresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.
- 50 **[0034]** Das erfindungsgemäße Zellkultursystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeföhrt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts föhren, sind dem Fachmann als adaptive

Mutationen bekannt. Die Erfindung umfaßt deshalb auch Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Sie umfaßt desweiteren ein Verfahren zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, sowie zellkultur-adaptierte Mutanten von HCV-RNA-Konstrukten, HCV-Vollängengenomen und HCV-Teilgenomen mit im Vergleich zu den ursprünglichen Konstrukten, Teil- oder Vollängengenomen erhöhter Replikationseffizienz.

**[0035]** Das erfindungsgemäße Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, ist dadurch gekennzeichnet, daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 19 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

**[0036]** Bei einer vorteilhaften Weiterbildung dieses Herstellungsverfahrens werden die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 eingeschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone geerntet und daraus die HCV-RNA-Konstrukte isoliert.

Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

**[0037]** Das erfindungsgemäße Verfahren zur Herstellung von Mutanten eines HCV - Vollängengenoms oder eines HCV- Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV- Vollängengenom oder - Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz zeichnet sich dadurch aus, daß man mit Hilfe eines der beiden vorstehend genannten Herstellungsverfahren eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt, diese aus den Zellen isoliert, mit im Stand der Technik bekannten Methoden kloniert und sequenziert und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Mutationen bestimmt, und diese Mutationen dann entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, welche die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängen- oder -teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Markergen die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

**[0038]** Die erfindungsgemäßen zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotid- und/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

**[0039]** Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nicht-selektierbaren Reportergen (z.B. Luziferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektionierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

**[0040]** Eine Gruppe ganz bevorzugter HCV-RNA-Konstrukte, HCV-Vollängengenome und HCV-Teilgenomen mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist

dadurch gekennzeichnet, daß sie einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der folgenden Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

5

**Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:**

**SEQ ID-NO: 1**

10 [0041]

Name: I389/Core-3'/wt  
Aufbau (Nukleotidpositionen):

15

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

20

4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7

25

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)

30

14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

**SEQ ID-NO: 2**

35 [0042]

Name: I337/NS2-3'/wt  
Aufbau (Nukleotidpositionen):

40

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

45

4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8407-8637: HCV 3' nicht-translatierte Region

55

## SEQ ID-NO: 3

[0043]

- 5 Name: I389/NS3-3'/wt  
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
  2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
  - 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
  4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
  5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
  - 15 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
  7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
  8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
  9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
  10. 7771-8001: HCV 3' nicht-translatierte Region

## SEQ ID-NO: 4

[0044]

- 25 Name: I337/NS3-3'/wt  
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
  2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
  - 30 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
  4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
  5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
  - 35 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
  7. 3859-4641: Nichtstrukturprotein 4B (NS4B)
  8. 4642-5982: Nichtstrukturprotein 5A (NS5A)
  9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
  10. 7759-7989: HCV 3' nicht-translatierte Region

## SEQ ID-NO: 5

[0045]

- 45 Name: I389/NS2-3'/wt  
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
  2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
  - 50 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
  4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
  5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
  - 55 6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
  7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
  8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
  9. 5302-6642: Nichtstrukturprotein 5A (NS5A)

10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase

11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

5

[0046]

Name: I389/NS3-3'/9-13F

Aufbau (Nukleotidpositionen):

10

1. 1-341: HCV 5' nicht-translatierte Region

2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

15

4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase

6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor

20

7. 3871-4653: Nichtstrukturprotein 4B (NS4B)

8. 4654-5994: Nichtstrukturprotein 5A (NS5A)

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

25

SEQ ID-NO: 7

[0047]

Name: I389/Core-3'/9-13F

Aufbau (Nukleotidpositionen):

30

1. 1-341: HCV 5' nicht-translatierte Region

2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

35

4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

5. 1813-2385: HCV Core Protein; Strukturprotein

6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein

40

7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein

8. 4051-4239: Protein p7

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease

10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase

11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor

45

12. 6946-7728: Nichtstrukturprotein 4B (NS4B)

13. 7729-9069: Nichtstrukturprotein 5A (NS5A)

14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase

15. 10846-11076: HCV 3' nicht-translatierte Region

50

SEQ ID-NO: 8

[0048]

Name: I389/NS3-3'/5.1

Aufbau (Nukleotidpositionen):

55

1. 1-341: HCV 5' nicht-translatierte Region

2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5B (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

## SEQ ID-NO: 9

[0049]

Name: I389/Core-3'/5.1

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

## SEQ ID-NO: 10

[0050]

Name: I389/NS3-3'/19

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 11

[0051]

- 5 Name: I389/Core-3'/19  
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
  2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
  - 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
  4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
  5. 1813-2385: HCV Core Protein; Strukturprotein
  - 15 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
  7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
  8. 4051-4239: Protein p7
  9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
  10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
  - 20 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
  12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
  13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
  14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
  15. 10846-11076: HCV 3' nicht-translatierte Region

25 [0052] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

30 **Fig. 1 A:** Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo<sup>R</sup>), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck ( $\Delta$ ) markiert.

45 **Fig. 1 B:** Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.

50 **Fig. 1 C** : Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone. Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.

55 **Fig. 2 A:** Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I<sub>377</sub>/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I<sub>377</sub>/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1  $\mu$ g 9-13 DNA vor der PCR (zwecks Ausschluß

eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem µg Gesamt-DNA des Zellkons 9-13 erhalten wurde.

- 5 **Fig. 2 B:** Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus" -Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrang" und "Plusstrang" bezeichnen die Spezifität der radioaktiven RNA-Sonden.
- 10 **Fig. 2 C:** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.
- Fig. 3 A:** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.
- 15 **Fig. 3 B:** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.
- 20 **Fig. 4:** Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.
- 25 **Fig. 5:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Antibiotikumresistenzgen (A) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3' NTR (bicistronische RNA).
- 30 **Fig. 6:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reporterprotein wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schraffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (D) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).
- 35 **Fig. 7:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.
- 40 **Fig. 8:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.
- 45 **Fig. 9:** Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in
- 50
- 55

naïve Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1 beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphorimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den in vitro Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 µg eines Plasmids zugesetzt, das die Expression der Luziferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luziferaseaktivität im jeweiligen Zellysat bestimmt. Die Anzahl der Kolonien wird jeweils auf die Luziferaseexpression normiert.

**Fig. 10:** Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukts I377/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den reklonierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* I verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.

**Fig. 11 A:** Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt I<sub>389</sub>/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luziferasegen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NS5B RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angedeutet. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine in vitro Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naïve Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.

**Fig. 11 B:** Vergleich der Luziferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt I<sub>389</sub>/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-13F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luziferaseaktivitäten luminometrisch bestimmt.

**Fig. 12:** Selektionierbare HCV-Vollängengenome (Konstrukte I<sub>389</sub>/core-3'/5.1 und I<sub>389</sub>/core-3'/9-13F).

(A) Schematische Darstellung des Vollängenkonstrukts. Der Bereich zwischen den beiden ange-deuteten Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hoch-adaptierten RNA-Varianten 5.1. oder 9-13F.

(B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 µg in vitro transkribierter RNA der unter A dargestellten Konstrukte I<sub>389</sub>/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experimentes.

(C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden in vitro Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das *neo*-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10<sup>8</sup> Molekülen der angegebenen in vitro Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 µg Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von in vitro transkribierter I<sub>389</sub>/core-3'/5.1-RNA bzw. I<sub>389</sub>/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

**Fig. 13:** HCV-RNA-Konstrukte mit einem Reportergen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reportergen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reportergenprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reportergenprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.

**Fig. 14:** Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.

**Fig. 15:** Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

### Beispiel 1: Herstellung von HCV-RNA-Konstrukten

#### (A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

**[0053]** Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h. die HCV-RNA wie nachfolgend beschrieben isoliert:

**[0054]** Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987,

Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGGCCTATTGG CCTGGAG) und dem 'expand reverse transcriptase'-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR=polymerase chain reaction) durchgeführt, und zwar unter Verwendung des 'expand long template'-Systems (Boehringer Mannheim,

Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde

1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde

1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACA-GCCCGCGTCATAGCACTCG) oder S4542 und A9386 (TTAGTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30

Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment

wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden

kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten

'X-tails' (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor inseriert. Abweichungen von der Konsensussequenz wurden mittels ortsgerichteter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3' Ende der 3' NTR (Kolykhalov et al. ibid.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTREN wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch

gefunden.

[0055] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

#### (B) Synthese selektierbarer HCV-Subgenom-Konstrukte

[0056] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, *Molecularcloning: a laboratory manual*, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, *Current Protocols in Molecular Biology*, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5' NTR inseriert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebenso gut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstrukts inseriert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Enzephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0057] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt: Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kotransfektionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0058] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I<sub>377</sub>/NS2-3' (oder I<sub>377</sub>/NS3-3') und I<sub>389</sub>/NS2-3' (oder I<sub>389</sub>/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

[0059] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I<sub>377</sub>/NS2-3' (oder I<sub>377</sub>/NS3-3') und I<sub>389</sub>/NS2-3' (oder I<sub>389</sub>/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0060] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

#### (C) Synthese selektierbarer HCV-Genom-Konstrukte

[0061] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasgens und der vollständigen EMCV-IRES verbunden ist, wurde mit NcoI und SpeI restringiert und mittels präparativer Agarosegelelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem NcoI/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/SpeI-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasgenfragment und der EMCV-IRES liegen, wurde danach mit PmeI und SpeI restringiert und mit dem analog restringierten I<sub>389</sub>/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

#### (D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Transkripten

[0062] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Transkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM MgCl<sub>2</sub>, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/µl, 50 µg/ml restringierte DNA und ca. 2 Einheiten/µl T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, *Bio Techniques* 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro µg DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittel optischer Dichtemessungen quantifiziert und

ihre Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

## **Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7**

5 [0063] Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription ( Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

10 [0064] Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, *Cancer Res.* 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in  $8 \times 10^6$  Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

15 [0065] Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klon 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.

20 [0066] Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

25 [0067] Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, *Anal. Biochem.* 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von  $0,5 - 1 \times 10^6$  Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig  $10^9$  in-vitro-Transkripte (ivtr.), die zu den  $I_{389}/NS2-3'/wt$  oder den  $I_{389}/NS3-3'/wt$  Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.

40 [0068] Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40 µg/ml, 1h, 37°C) in 10mM Tris, pH7,5, 1mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend  $4 - 8 \times 10^4$  Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTTCCGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit  $10^7$  Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.

Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrol-

len, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

**[0069]** Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren:

Es wurden jeweils  $10^6 - 10^2$  Plasmid Moleküle ( $I_{377}/NS3-3'/wt$ ) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils  $1 \mu g$  9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem  $\mu g$  DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder  $2 \mu g$  Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrang', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrang', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca.  $10^8$  Kopien/ $\mu g$  Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [ $^3H$ ] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [ $^3H$ ] sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca.  $5 \times 10^5$  Zellen der Klone 9-13 und 8-1 mit  $100 \mu Ci$  [ $^3H$ ]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von  $4 \mu g/ml$  Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 und Lohmann et al., 1997, *J Virol.* 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflusst worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Reklonierung der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der reklonierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transkribierten Transkript des HCV-Konstrukts  $I_{377}/NS3-3'/wt$  übereinstimmt.

**[0070]** Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [ $^{35}S$ ] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 ( $\Delta$ ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nicht-denaturierenden Bedingungen (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519) und unter Verwendung von drei verschiedenen Antiseren (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abge-

trennt worden.. Die Immunkomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonkonstrukt I<sub>377</sub>/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum (3/4') bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0071] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transients Expression desselben bicistronischen HCV-RNA-Konstrukts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunfluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antisera durchgeführt (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 ( $\Delta$ ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azeton fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antisera inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegengefärbt.

[0072] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antisera war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antisera führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen. Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet — so wie das bei den meisten RNA-Viren der Fall ist.

[0073] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellsystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

### **Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten**

[0074] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektioniert, wobei hier jedoch selektionierbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [<sup>3</sup>H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunpräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

### **Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukts, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.**

[0075] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das inklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende

Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

#### **Beispiel 5: Kotransfektion eines HCV-RNA-Konstrukts mit einem separaten Luziferase-Transfektionskonstrukt**

[0076] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen vermittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen, insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen. Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt, die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestimmung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann ebenso ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmarkers ein Maß für die Replikation der viralen DNA.

#### **Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfähren für die Gentherapie**

[0077] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementierende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispielsweise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also virus-ähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des inklonierten Fremdgens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Konstrukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäße HCV-Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vorgenannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als Expressionsvektor.

#### **Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte**

##### **(A) Isolationsverfahren**

[0078] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca.  $10^9$  Moleküle HCV-RNA enthielt, wurde anschließend per Elektroporation in naive Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden  $10^9$  in vitro Transkripte der analogen neo-HCV-RNA, die mit isolierter Gesamt-RNA aus naiven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg aufgefüllt worden war, in naive Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien, ausgedrückt in 'colony forming units (cfu) pro Mikrogramm RNA' in den beiden Ansätzen bestimmt. Bei einer Konzentration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der *Isolierten* Gesamt-RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit dersel-

ben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

**[0079]** Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplifikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser reklonierten RNAs mit der Sequenz der RNA, die ursprünglich in die naiven Huh-7 Zellen eingeschleust wurde ergab, daß die reklonierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10). *Sfi*I-Fragmente dieser reklonierten Mutanten wurden im Austausch gegen das analoge *Sfi*I-Fragment des ursprünglichen Replikonkonstrukts in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naive Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt. Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der reklonierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten. Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.

**[0080]** Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naive Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.

Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den *neo*-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektionieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.

Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naiven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt I<sub>389</sub>/NS3-3' erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IRES besaß (I<sub>377</sub>/NS3-3'; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naive Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naive Huh-7 Zellen transfiziert und analog selektioniert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die *neo*-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfi*I verdaut und in das *Sfi*I-restringierte Ausgangskonstrukt I<sub>389</sub>/NS3-3' inseriert. Insgesamt wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naive Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten *neo*-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. "adaptiven Mutationen") mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNA.

## (B) Modifikationsverfahren

**[0081]** Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukts. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht

werden können, die kein selektierbares Markergen mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffizienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprechen. Zwecks einfacher Messung wurde das *neo*-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NS5B RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

#### Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome

[0082] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, der die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das Sfil-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transferiert (Fig. 12). Dieses HCV-Genom wurde in naive Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NS5B RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängengenom enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierten HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfindungsgemäßen System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfindungsgemäßen HCV-Zellkultursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

#### Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.

[0083] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

**Beispiel 10:** Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfähre für die Gentherapie oder als Expressionsvektor.

[0084] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsidiert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

**Beispiel 11:** Herstellung von monoclonalen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0085] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Tabelle 1

Spezifische Infektiositäten (cfu/µg RNA) der HCV-RNA-Konstrukte mit adaptiven Mutationen, die bei der 9-13F Mutante gefunden und in das parentale HCV-RNA-Konstrukt I <sub>389</sub> /NS3-3'/wt eingeführt wurden		
Aminosäureaustausch <sup>1</sup>	HCV-Protein	cfu/µg RNA <sup>2</sup>
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

<sup>1</sup> Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

<sup>2</sup> Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500µg/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/ $\mu$ g RNA) des parentalen HCV-RNA-Konstrukts I <sub>388</sub> /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte RNA-Variante	cfu/ $\mu$ g RNA <sup>1</sup>	
Wildtyp	30 - 50	
9-13 C	100 - 1.000	
9-13 I	100 - 1.000	
9-13 F	1.000 - 10.000	
5.1	50.000 - 100.000	
19	50.000 - 100.000	

<sup>1</sup> Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 $\mu$ g/ml G418.

**Tabelle 3:** Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-RNA-Konstrukt I<sub>389</sub>/NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19

HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
9-13 F	9346	T > C	Val > Ala
	3866	C > T	-
	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
5.1	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
	8486	C > T	-
	8991	C > G	Arg > Gly
	4180	C > T	Thr > Ile
	4679	C > T	-

5		4682	T > C	-
		5610	C > A	Leu > Ile
		6437	A > G	-
10		6666	A > G	Asn > Asp
		6842	C > T	-
		6926	C > T	-
15		6930	T > C	Ser > Pro
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
20	NK19	3946	A > G	Glu > Gly
		4078	C > G	Ala > Gly
		4180	C > T	Thr > Ile
25		4682	T > C	-
		5610	C > A	Leu > Ile
30		5958	A > T	Met > Leu
		6170	T > A	-
		6596	G > A	-
35		6598	C > G	Ala > Gly
		6833	C > T	-
		6842	C > T	-
40		6930	T > C	Ser > Pro
		7141	A > G	Glu > Gly
		7320	C > T	Pro > Ser
45		7389	A > G	Lys > Glu
		7735	G > A	Ser > Asn

50

[0086] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen der zellkulturadaptierten HCV-RNAs. Die Zahlen  
55 beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

## SEQUENCE LISTING

<110> Bartenschlager, Ralf

<120> Hepatitis C Virus cell culture system

<130> ba-1

<140> 199 15 178.4

<141> 1999-04-03

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 11076

<212> DNA

<213> Hepatitis C Virus

<400> 1

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggtctg cggaaccggt gactacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgttact gcctgatagg 300
gtgcttgcyg gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ceggttcttt 540
ttgtcaagac cgacctgtcc ggtgcccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaaatgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggtacactg ccatttcgac caccagcgca aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgcgat caggatgatc tggacgaaga gcacagggg ctcgcgccag 900
ccgaactgtt cgcagggctc aaggcgcgca tccccgacgg caggatcttc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggtcg accgcttctc cgtgttttac ggtatcgccg 1140
ctccgcattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagttaaa 1200
cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg ctggaataaa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatag tgcctgtttt ttgcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtgtc ttcccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaaccccc accctggcgac aggtgcctct gcggccaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcaaacccc cagtgcacag ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatg gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgagggtta aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccattggcac gaatcctaaa cctcaaagaa aaaccaaag taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcggt gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggccttg 2040
gctcagcccc ggtacccctg gccctctctat ggcaatgagg gcttgggtgt ggacaggatg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cgacccccg gcgtagggtc 2160
cgcaatttgg gtaaggctcat cgataccctc acgtgcggct tcgcccgtct catgggttac 2220
attccgctcg tcggcgcccc cctagggggc gctgccagg ccctggcgca tggcgctccg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttctc cttttctatc 2340
ttccttttgg ctttctgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag gttaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccggtgtc gtgccctgcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt gggggcggtg ctctctgtc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgccagc tgttcacct ctcgcctcgc 2700

```

5

10

15

20

25

30

35

40

45

50

55

cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760  
 cgtatggcct gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820  
 ttactccgga tcccacaagc tgctgtggat atggtggcgg gggccctatg gggagtccta 2880  
 gcgggccttg cctactatcc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940  
 ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggcagatggc caaaaacacc 3000  
 ctccgggatta cgtccctctt ttcacccggg tcatcccaga aaatccagct tgtaaacacc 3060  
 aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120  
 ttctttgctg cgtgtttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180  
 gccagctgca gcccacatcga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240  
 tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtgcggtatc 3300  
 gtacccgcgg cgcaggtgtg tggtccagtg tactgttca cccaagccc tgcgtgggtg 3360  
 gggacgacgg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420  
 ctgcttctta acaacacggg gccgcccga ggaactggg ttggtgtgac atggatgaat 3480  
 agcactgggt tcaccaagac gtgcgggggc ccccggtgta acatcggggg gatcggaat 3540  
 aaaaccttga cctgccccac ggactgcttc cggaaagcacc ccgagggcac ttacaccaag 3600  
 tgtggttcgg gccttgggtt gacacccaga tgcttggctc actaccata caggccttgg 3660  
 cactaccctt gcactgtcaa ctttaccatc ttcaaggta ggatgtacgt ggggggagtg 3720  
 gagcacaggc tcgaagccgg actcgaggag agcgttgtaa cctggaggag cctggaggag 3780  
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtggca ggtattgccc 3840  
 tgttctctca ccacctacc ggctctgtcc actggttga tccatctcca tcagaacgtc 3900  
 gtggacgtac aatacctgta cggatatagg tcggcggttg tctcctttgc aatcaaatgg 3960  
 gagtatgtcc tgtgtctctt ccttctcttg gcggacggcg gcgtctgtgc ctgcttggg 4020  
 atgatgtgct tgatagctca agctgaggcc gccctagaga acctggtggt cctcaacggc 4080  
 gcatccgtgg ccggggcgca tggcaattct tcttctctcg tgttctctg tgcgtccctg 4140  
 tacatcaagg gcaggctggc ccttggggcg gcatatgccc tctacggcgt atggccgcta 4200  
 ctctctgctc tgctggcggtt accaccacga gcatacggca tggaccggga gatggcagca 4260  
 tcgctcggag gcgcggtttt cgtaggtctg atactcttga ccttgtcacc gcactataag 4320  
 ctgttctctg ctaggctcat atggtggtta caatatatta tcaccaggcg cagggcacac 4380  
 ttgcaagtgt ggateccccc cctcaacgct cggggggggc gcgatgccgt catcctcctc 4440  
 acgtgcggca tccaccacga gctaactctt accatcacca aaatcttgc ctgcatactc 4500  
 ggtccactca tgggtgtcca ggctggtata accaaagtgc cgtactctgt gcgcgcacac 4560  
 gggctcatcc gtgcatgcat gctggtgctg aaggttgcgt ggggtcatta tgtccaatg 4620  
 gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccocactg 4680  
 cgggactggg cccacggggg cctacgagac ctgcggtgg cagttgagcc cgtcgtcttc 4740  
 tctgatattg agaccaaggt tatcacctgg ggggcagaca ccgcgccgtg tggggacatc 4800  
 atcttggggc tgcccgctct cgcgcgaggg gggagggaga tacatctggg accggcagac 4860  
 agccttgaag ggcagggggt ggcactctc gcgcctatta cggcctactc ccaacagacg 4920  
 cgaggcctac ttggctgcat catcactagc ctcacagggc gggacaggaa ccaggtcgag 4980  
 ggggaggctc aagtgtctct caccgcaaca caatcttctc tggcgacctg cgtcaatggc 5040  
 gtgtgttggg ctgtctatca tgggtgccgg tc aaagaccc ttgccggccc aaagggccca 5100  
 atcacccaaa tgtacacca tggtagccag gacctcgtcg gctggcaagc gccccccggg 5160  
 gcgcgttctt tgacaccatg cacctgcggc agctcggacc tttacttggg caccaggcat 5220  
 gccgatgtca ttccggtgcg ccggcggggg gacagcaggg ggagcctact ctcccccagg 5280  
 cccgtctctc acttgaaggc ctcttcgggc ggtccactgc tctgccctc ggggcacgtc 5340  
 gtgggcatct ttccggctgc cgtgtgcacc cagggggttg cgaaggcggt ggactttgta 5400  
 cccgtcgagt ctatgaaac cactatgcgg tccccggtct tcacggacaa ctgcgtccct 5460  
 ccggccgtac cgcagacatt ccagggtggc catctacacg ccctactgg tagcggcaag 5520  
 agcactaagg tgccggtgct gtagcagcc caagggtata aggtgcttgt cctgaacccc 5580  
 tccgtcgccg ccacctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640  
 aacatcagaa ccggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700  
 ggcaagtttc ttgccgacgg tggttgctct gggggcgcc atgacatcat aatattgat 5760  
 gagtgccact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820  
 gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggtcacc 5880  
 gtgccacatc caaacatcga gagggtggct ctgtccagca ctggagaaat ccccttttat 5940  
 ggcaaaagcca tccccatcga gaccatcaag ggggggaggg acctcatttt ctgccattcc 6000  
 aagaagaaat gtgatgagct cgcgcggaag ctgtccggcc tggactcaa tgcgttagca 6060  
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagcgtcat tgcgttagca 6120  
 acggacgctc taatgacggg ctttaccggc gatctcgact cagtgatcga ctgcaatata 6180  
 tgtgtcaccg agacagtcga cttcagcctg gacccgacct tcaccattga gacgacgacc 6240  
 gtgccacaag acgcggtgtc acgctcgag cggcgaggca ggactggtag gggcaggtg 6300  
 ggcatattaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttcccggtt 6360  
 ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420  
 gttaggttgc gggcttacct aaacacacca ggggtgccc tctgccagga ccatctggag 6480  
 ttctgggaga gcgtctttac aggcctcacc ccaatttctt gtcccagact 6540  
 aagcaggcag gagacaactt cccctacctg gttagcatacc aggtacgggt gtgcgccagg 6600  
 gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatagc gctaaagcct 6660

	acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
	accacacacc	ccataacca	atcacatcatg	gcatgcatgt	cggtcgacct	ggaggtcgtc	6780
	acgagcacct	gggtgctggt	aggcgagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
5	acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattecc	6900
	gacagggaa	tcctttaccg	ggagttcgat	gagatggaag	agtgcgcctc	acacctccct	6960
	tacatcgaa	agggaaatgca	gctcgcgcaa	caattcaaac	agaaggcaat	cgggttgcgtg	7020
	caaacaagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaatccaa	gtggcggacc	7080
	ctcgaagcct	tctgggagaa	gcatatgtgg	aatttcatca	gcgggataca	atatttagca	7140
	ggcttgtcca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
	atcaccagcc	cgcctaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
10	gcccgcaca	ttgtcctctc	cagcgtgctc	tctgctttcg	taggcgcggg	catcgctgga	7320
	gcggctgttg	gcagcatagg	ccttgggaag	gtgcttggg	atattttggc	aggttatgga	7380
	gcaggggtgg	caggcgcgct	cgtggccttt	aaggtcatga	gcggcgagat	gccctccacc	7440
	gaggacctgg	ttacctaact	ccctgtctatc	ctctcccctg	gcgccctagt	cgtcgggggtc	7500
	gtgtgcgcag	cgatactcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcgggg	aaccacgtct	ccccacgcga	ctatgtgcct	7620
	gagagcgagc	ctgcagcacg	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
15	ctgaagaggc	ttcaccagt	gatcaacgag	gactgtctca	cgccatgctc	cggtcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	ttcaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaacat	tgcccattaa	cgcgtacacc	acgggcccct	gcacgccttc	cccggcgcca	8040
20	aattattcta	gggcgctgtg	gcgggtggct	gctgaggagt	acgtggaggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtggcc	gtgtcaggtt	8160
	ccggcccccg	aattcttcac	agaagtggat	ggggtgcggt	tgacaggta	cgctccagcg	8220
	tgcaaacccc	ttctacggga	ggaggtcaca	ttcctggtcg	ggctcaatca	atacctggtt	8280
	gggtcacagc	ttccatgcga	gcccgaaccg	gacgtagcag	tgctcacttc	catgctcacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
	tccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
25	accgctcatg	actccccgga	cgtgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atggggcgga	acatcacccg	cgtggagtca	gaaaataag	tagtaatttt	ggactctttc	8580
	gagcgcctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
	aggteccagg	aattccctcg	agcgtgccc	atatgggcac	gcccggatta	caacccctca	8700
	ctggttagagt	cctggaagga	cccgagctac	gtcccctcag	tggatcacag	gtgtccattg	8760
	ccgcctgcc	aggccccctc	gataccacct	ccacggagga	agaggacggt	tgtcctgtca	8820
30	gaattctacc	tgtcttctgc	cttggcgag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tcgtcgcccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggctt	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgctgct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgcggag	9120
	gaaaccaagc	tgcccatcaa	tgactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
	tatgtctaca	catctcgacg	cgaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
35	caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcgccc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgcgt	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	agggggggcg	caagccagct	9540
	cgccttatcg	tattcccaga	tttgggggtt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
40	gtggtctcca	ccctccctca	ggcctgatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatggcc	tatgggcttc	9720
	gcatatgaca	cccgtgtttt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
	tcaatctacc	aatgtttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtctaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cggtgcgcgg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
45	aaggccgctg	cggcctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
	gacgaccttg	tcgttatctg	tgaagccgcg	gggacccaag	aggacgaggg	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
	ggcaaaaagg	tgtactatct	cacccgtgac	cccaccaccc	cccttgcgcg	qgctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
50	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
	tctccagggtg	agatcaatag	ggtggcttca	tgcctcagga	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgtagggc	tactgtccca	gggggggagg	10620

gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680  
ccaatcccgg ctgctgcca gttggattta tccagctggt tcgttgctgg ttacagcggg 10740  
ggagacatat atcacagcct gtctcgtgcc cgaccccgct ggttcatgtg gtgcctactc 10800  
5 ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860  
ctccaggcca ataggccatc ctgttttttt cccttttttt ttttcttttt tttttttttt 10920  
tttttttttt tttttttttc tecttttttt ttectctttt tttccttttc tttccttttg 10980  
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040  
agagagtgct gatactggcc tctctgcaga tcaagt 11076

10

15

20

25

30

35

40

45

50

55

<210> 2  
 <211> 8637  
 <212> DNA  
 <213> Hepatitis C Virus

5

```

<400> 2
gccagccccc gattgggggc gacactccac catagatcac tccccgtgga ggaactactg 60
tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgctcgtcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
gacgacgggy tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgttact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggcccgctt ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttccgg ctgtcagcgc agggggcgcc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacggggcg tccttgccga gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctatttgg cgaagtgcgg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaaa catcgcatcg agcagacacg tactcggatg gaagccggtc 840
ttgtcgatca ggtgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900
ccaggctcaa ggcgcgcag cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960
gcttgccgaa tatcatggtg gaaaatggcc gcttttcttg attcatcgac tgtggccggc 1020
tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080
ttggcggcga atgggctgac cgttctctcg tgctttacgg tatcgccgct cccgattcgc 1140
agcgcacatc cttctatcgc cttcttgacg agttcttctg agtttaaacg gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tggataaagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcgggaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt cccctctcgc caaagggaat caaggtctgt tgaatgtcgt gaaggaagca 1440
gttcccttgg aagcttcttg aagacaaaca acgtctgtag cgacctttg caggcagcgg 1500
gaccccccac ctggcgacag gtgcctctgc ggccaaaagc cactgtgata agatacacct 1560
gcaaaggcgg cacaacccca gtgccacggt gtgagttgga tagttgtgga aagagtcaaa 1620
tggtctctct caagcgtatt caacaagggg ctgaaggatg ccagaaaggc accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1740
aacgtctagg cccccgaac cacggggacg tggtttctct ttgaaaaaca cgataatacc 1800
atggaccggg agatggcagc atcgtgcgga ggcgcgggtt tcgtaggctc gatactcttg 1860
accttgtcac cgcactataa gctgttctct tatgggtggtt acaatatttt 1920
atcaccaggg ccgaggcaca cttgcaagtg tggatcccc cctcaacgt tcggggggggc 1980
cgcgatgcgg tcactctcct cacgtgcgcg atccaccag agctaactct taccatcacc 2040
aaaatcttgc tcgcataact cgttccactc atggtgtctc aggcctggtat aaccaaagtg 2100
ccgtactctg tgccgcgaca cgggctcatt cgtgcatgca tgctggtgcg gaaggttgct 2160
gggggtcatt atgtccaaat ggtctctcag aagttggcgg cactgacagg tacgtacgtt 2220
gatgaccatc tcaccccatc gcgggactgg gagaccaagg ttatcacctg gggggcagac 2280
gcagttgagc ccgtctctct ctctgatag gagaccaagg ctgcccgtct ccgcccgcag ggggagggag 2340
accgcgcggt gtggggacat catcttgggc ctgcccgtct ccatcactag cctcacaggc 2400
atacatctgg gaccggcaga cagccttgaa gggcaggggt ggcgactcct cgcgcctatt 2460
acggcctact cccaacagac gcgaggccta cttggctgca tcactactag cctcacaggc 2520
cgggacagga accaggtcga gggggaggtc caagtggctc ccaccgcaac acaatcttct 2580
ctggcgacct gcgtcaatgg cgtgtgttgg actgtctatc atggtgccgg ctcaaagacc 2640
cttgccggcc caaaggggcc aatcacccaa atgtacacca atgtggacca ggacctcgtc 2700
ggctggcaag cgcctcccg ggcgcgttcc ttgacaccat gcacctgcgg cagctcggac 2760
ctttacttgg tcacgaggca tgccgatgtc attccggtgc gccggcgggg cgacagcagg 2820
gggagcctac tctccccag gccgctctcc tacttgaagg gctcttcggg cgttccactg 2880
ctctgcccc ctggggcacgc tgtgggcac tttcgggtcg ccgtgtgcac ccgaggggtt 2940
gcgaaggcgg tggacttgt acccgtcgag tctatggaaa ccaatagtcg gtccccggtc 3000
ttcacggaca actcgtcccc tccggccgta ccgcagacat tccagggtgg ccatctacac 3060
gccctactg gtagcggcaa gagcactaag gtgccggctg cgtatgcagc ccaagggtat 3120
aaggtgcttg tcctgaaccc gtccgtcgcc gccaccctag gtttcggggc gtatatgtct 3180
aaggcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac cacgggtgcc 3240
ccatcacgt actccacctg tggcaagttt cttgcccagc gtggttgcct tgggggcgcc 3300
tatgacatca taatatgtga tgagtggcc tcaactgact cgaccactat cctgggcac 3360
ggcacagtcc tggaccaagc ggagacggct ggagcgcgac tcgtcgtgct cgcaccgct 3420
acgcctccgg gatcggtcac cgtgccacat ccaaacatcg aggaggtggc tctgtccagc 3480
actggagaaa tcccccttta tggcaaaagc atccccatcg agaccatcaa gggggggagg 3540

```

55

5

10

15

20

25

30

35

40

45

50

55

```

cacctcattt tctgccattc caagaagaaa tgtgatgagc tcgccgcgaa gctgtccggc 3600
ctcggactca atgctgtagc atattaccgg ggccttgatg tatccgctcat accaactagc 3660
ggagacgtca ttgtcgtagc aacggacgct ctaatgacgg gctttaccgg cgatttcgac 3720
tcagtgtatc actgcaatac atgtgtcacc cagacagtcg acttcagcct ggaccgcgac 3780
ttcaccattg agacgacgac cgtgccacaa gacgcggtgt caccgtcgca gcggcgaggc 3840
aggactggta ggggcaggat gggcatttac aggtttgtga ctccaggaga acggccctcg 3900
ggcaatgttc attcctcggg tctgtgcgag tgctatgacg cgggctgtgc ttgggtacgag 3960
ctcacgcgcg ccgagacctc agtttaggtg cgggcttacc taaacacacc agggttgccc 4020
gtctgccagg accatctgga gttctgggag agcgtcttta caggcctcac ccacatagac 4080
gcccatttct tgtccagac taagcaggca ggagacaact tcccctacct ggtagcatac 4140
caggctacgc tgtgcgccag ggctcaggct ccacctccat cgtgggacca aatgtggaag 4200
tgtctcatat ggctaaagcc tacgtgcac gggccaacgc ccctgctgta taggctggga 4260
gccgttcaaa acgaggttac taccacacac ccataacca aatacatcat ggcattgcatg 4320
tcggctgacc tggaggtcgt cagcagcacc tgggtgctgg taggcggagt cctagcagct 4380
ctggccgctg attgctgac aacaggcagc gtggtcattg tgggcaggat catcttgtcc 4440
ggaaagcccg ccatcattcc cgacagggaa gtcccttacc gggagtctga tgagatggaa 4500
gagtgcgcct cacacctccc ttacatcgaa cagggaatgc agctcgccga acaattcaaa 4560
cagaaggcaa tcgggttgct gcaaacagcc accaagcaag cggaggtgct tgctcccgctg 4620
gtggaatcca agtggcgagc cctcgaagcc ttctggcgga agcatatgtg gaatttcac 4680
agcgggatac aatatttagc aggtctgtcc actctgcctg gcaaccccg ccatagcatca 4740
ctgatggcat tcacagcctc tatcacagc cgcctacca ccaacatac cctcctgttt 4800
aacatcctgg ggggatgggt ggccgcccaa ctgtctctc ccagcgtgc ttctgtcttc 4860
gtaggcgccg gcatcgctgg agcggctgtt ggcagcatag gecttgggaa ggtgcttgtg 4920
gatatttttg cagggttatg agcaggggtg gcaggcgcg cctgtggcctt taaggctag 4980
agcggcgaga tgccctccac cgaggacctg gtaacctac tccctgctat cctctccctc 5040
ggcgccctag tcgtcggggt cgtgtgcgca gcgatactgc gtcggcacgt gggcccgagg 5100
gagggggctg tgcagtggat gaaccggctg atagcgttcg cttcgcgggg taaccacgtc 5160
tccccacgc actatgtgct tgagagcgac gctgcagcac gtgtactca gatcctctct 5220
agtcttacca tcaactagct gctgaagagg ctccaccagt ggatcaacga ggactgtccc 5280
acgccatgct ccggctcgtg gctaagagat gtttgggatt ggatatgcac ggtgttgact 5340
gatttcaaga cctggctcca ctgcccgat tcgggggagt tgcggggagt cccctctctc 5400
tcatgtcaac gtgggtacaa gggagtctgg cggggcgagc gcatcatgca aaccacctgc 5460
ccatgtggag cacagatcac cggacatgtg aaaaacggtt ccatgaggat cgtggggcct 5520
aggacctgta gtaacacgtg gcatggaaac tccccatta cagcgtagac caggggcccc 5580
tgcaagccct ccccgcgccc aaattattct agggcgctgt ggcgggtggc tgctgaggag 5640
tacgtggagg ttacgcgggt gggggatttc cactacgtga cgggcatgac cactgacaac 5700
gtaaagtgcg cgtgtcaggt tccggccccc gaattcttca cagaagtgga tgggggtcgg 5760
ttgcacaggt acgctccagc gtgcaaaccc ctccctacggg agggaggtcac attcctggct 5820
gggtccaatc aatacctggt tgggtcacag ctcccctgag agcccgaaac ggacgtagca 5880
gtgctcaatt ccattgtcac cgaccctcc cacttacgg agtgcagct ctagccagct 5940
ctggccaggg gatctcccc ctccctggcc agctcatcag ctgcccagct gtctgcgct 6000
tccttgaagg caacatgcac taccctgcat gactcccccg acgctgacct catcgaggcc 6060
aacctcctgt ggcggcagga gatgggcggg aacatcacc gcgtggagtc agaaaataag 6120
gtagttaatt tggactcttt cgagccgctc caagcggagg aggatgagag ggaagtatcc 6180
gttccggcgg agatcctgag gaggtccagg aaattccctc gagcgtatgc catatgggca 6240
cgcccggtat acaacctccc actgttagag tectggaagg acccggacta cgtccctcca 6300
gtggtacacg ggtgtccatt gccgcctgcc aaggccctc cgataccacc tccacggagg 6360
aagaggacgg ttgtcctgtc agaactctac gtgtcttctg ccttggcgga gctcgccaca 6420
aagaccttcg gcagctccga atcgtcgccc gtcgacagcg gcacggcaac ggcctctcct 6480
gaccagccct ccgacgacgg cgacgcggga tccgacgttg agtcgtact ctccatgcc 6540
ccccttgagg gggagccggg ggtatccgat ctacgcgag ggtcttggtc taccgtaagc 6600
gaggaggcta gtgaggcgt cgtctgctgc tcgatgtcct acacatggac aggcgcccgt 6660
atcacgcat gcgctcgga ggaaccaaag ctgcccata atgcaactg caactctttg 6720
ctccgtcacc acaacttgg ctatgctaca acatctcgca gcgcaagcct gcggcagaag 6780
aaggctacat ttgacagact gcaggtcctg gacgaccact accgggacgt gctcaaggag 6840
atgaaggcga aggcgtccac agttaaggct aaactctat ccgtggagga agcctgtaag 6900
ctgacgcccc cacttcggc cagatctaaa tttggtatg gggcaaggga cgtccggaac 6960
ctatccagca aggcctgtaa ccacatccgc tccgtgtgga aggacttgct ggaagacact 7020
gagacaccaa ttgacaccac catcatggca aaaaatgagg tttctcgtt ccaaccagag 7080
aagggggggc gcaagccagc tcgccttacc gtattccag atttgggggt tcgtgtgtgc 7140
gagaaaaatg ccctttacga tgtggtctcc accctccctc aggcctgat gggctcttca 7200
tacggattcc aatactctcc tggacagcgg gtcgagttcc tggatgaatg ctggaaagcg 7260
aagaaatgcc ctatgggctt cgcatatgac acccgctgtt ttgactcaac ggtcactgag 7320
aatgacatcc gtgttgagga gtaactctac caatgttgtg acttggcccc cgaagccaga 7380
caggccataa ggtcgtcac agagcggctt tacatcgggg gccccctgac taattctaaa 7440
gggcagaact gcggctatcg ccggtgccgc gcgagcggtg tactgacgac cagctgcggt 7500

```

5 aataccctca catgttactt gaaggccgct gcggcctgtc gagctgcgaa gctccaggac 7560  
 tgcacgatgc tcgtatgcgg agacgacctt gtcgttatct gtgaaagcgc ggggacccaa 7620  
 gaggacgagg cgagcctacg ggccttcacg gaggtatga ctagatactc tgccccccct 7680  
 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgtc ctccaatgtg 7740  
 tcagtcgcgc acgatgcac tgga aaaagg gtgtactatc tcaccogtga cccaccacc 7800  
 ccccttgccg gggctgcgtg ggagacagct agacacactc cagtcaattc ctggctaggc 7860  
 aacatcatca tgtatgcgcc caccttggtg gcaaggatga tcctgatgac tcatttcttc 7920  
 tccatccttc tagctcagga acaacttgaa aaagccctag attgtcagat ctacggggcc 7980  
 10 tgttactcca ttgagccact tgacctacct cagatcattc aacgactcca tggccttagc 8040  
 gcattttcac tccatagtta ctctccagggt gagatcaata ggggtggcttc atgcctcagg 8100  
 aaacttgggg taccgccctt gcgagtctgg agacatcggg ccagaagtgt ccgcgctagg 8160  
 ctactgtccc agggggggag ggctgccact tgtggcaagt acctcttcaa ctgggcagta 8220  
 aggaccaagc tcaaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280  
 ttcgttgctg gttacagcgg gggagacata tatcacagcc tgtctcgtgc ccgacccgc 8340  
 15 tggttcatgt ggtgcctact cctactttct gtaggggtag gcattctatct actccccaac 8400  
 cgatgaacgg ggagctaaac actccaggcc aataggccat cctgtttttt tccctttttt 8460  
 ttttctttt tttttttttt tttttttttt tttttttttt ctctttttt tttctctttt 8520  
 ttttctttt ctttctttt gtggctccat cttagcccta gtcacggcta gctgtgaaag 8580  
 gtccgtgagc cgcttgactg cagagagtgc tgatactggc ctctctgcag atcaagt 8637  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

<210> 3  
 <211> 8001  
 <212> DNA  
 <213> Hepatitis C Virus

5

<400> 3  
 gccagccccc gattgggggc gacactccac catagatcac tccctgtga ggaactactg 60  
 tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgcgtgcag cctccaggac 120  
 cccccctccc gggagagcca tagtggtctg cggaaacggg gagtacaccg gaattgccag 180  
 gacgaccggg tcccttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240  
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300  
 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360  
 ctcaaaagaaa aaccaaaccgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420  
 acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480  
 caatcggtcg gctgctattg ggcgaagtgc ggctgtcagc gcagggggcg ccggttcttt 540  
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600  
 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgc actgaagcgg 660  
 gaagggaact gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720  
 ctccctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780  
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcga 840  
 tgggaagcgg ccttgctgat caggatgac tggacgaaga gcatcagggg ctgcgcggac 900  
 ccgaactgtt cgccaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgaccc 960  
 atggcgatgc ctgcttgccg aatatcatgg tggaaatgg ccgcttttct ggattcatcg 1020  
 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080  
 ttgctgaaga gcttgccggc gaatgggctg accgcttctc cgtgctttac gctatcgccg 1140  
 ctcccgatcc gcagcgcac gccttctatc gccttcttga cgagttcttc tgagttttaa 1200  
 cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260  
 aacgttactg gccgaagcgg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320  
 tccaccatat tgccttcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380  
 acgagcatcc ctagggtgtt ttccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440  
 gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgacctgt 1500  
 tgcaggcagc gcaaccccc accctggcgc aggtgcctct gcggccaaaa gccacgtgta 1560  
 taagatacac ctgcaaaagg ggcacaaccc cagtgcacag ttgtgagttg gatagttgtg 1620  
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680  
 gtaccaccat gtatgggatc tgatctgggg cctcgttgca catgctttac atgtgtttag 1740  
 tgcaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800  
 cagcataata ccattggccc tattacggcc tactcccaac agacgcgagg cctacttggc 1860  
 tgcatcatca ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg 1920  
 gtctccacgg caacacaatc ttctctggcg acctgcgtca atggcggtg ttggactgtc 1980  
 tatcatggtg ccggctcaaa gaccttgcg ggcccaagg gcccaatcac ccaaatgtac 2040  
 accaatgtgg accaggacct cgtcggctgg caagcgcccc ccggggcgcg ttcttgaca 2100  
 ccactgcacct gcggcagctc ttggctcacg ggcattgccg ggcattccg tgtcattccg 2160  
 gtgcgcccgg ggggcgacag cagggggagc ctactctccc ccaggcccg ctctacttg 2220  
 aagggctctt cgggcggtcc actgctctgc ccctcggggc acgctgtggg catctttcgg 2280  
 gctgcgtgtg gcaaccgagg ggttgcgaag gcggtggact ttgtacccgt cgagtctatg 2340  
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400  
 acattccagg tggcccatct acacgcccc actggtagcg gcaagagcac taagggtgccc 2460  
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgtccgt cgccgcccac 2520  
 ctagggttcc gggcgatat gtctaaggca catggtatcg accctaaccat cagaaccggg 2580  
 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gttctctgcc 2640  
 gacggtggtt gctctgggg cgccatgac atcataatat gtgatgagtg cactcaact 2700  
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760  
 cgactcgtcg tgctcgccac cgctacgct ccgggatcgg tcaccgtgcc acatccaaac 2820  
 atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880  
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940  
 gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000  
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaata 3060  
 acgggcttta ccggcgattt cgactcagt atcgactgca atacatgtgt caccagaca 3120  
 gtcgacttca gcttggaacc gaccttcacc attgagacga cgaccgtgcc acaagacggc 3180  
 gtgtcacgct cgacggggcg aggcaggact ggtaggggca ggtagggcat ttacaggttt 3240  
 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300  
 gacgggggct gtgcttggtc cgagctcacg cccgcggaga cctcagttag gttgcgggct 3360  
 tacctaataa caccaggggt gcccgctcgc caggaccatc tggagttctg ggagagcgtc 3420  
 tttaacaggc tcaccacat agacgcccct ttctgtccc agactaagca ggcaggagac 3480  
 aacttcccct acctggtagc ataccaggct acgggtgtgc ccagggtcca ggctccacct 3540

55

5

10

15

20

25

30

35

40

45

50

55

```

ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacggggcca 3600
acgcccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acaccccata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtaggcg gagtccatgc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
atttgtggga ggatcatctt gtcgggaaag ccggccatca tccccgacag ggaagtcctt 3840
taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caagcggagg ctgctgctcc cgtggtggaa tccaagtggc ggaccctcga agccttcttg 4020
gcgaagcata tgtggaatth catcagcggg atacaatatt tagcaggctt gtccactctg 4080
cctggcaacc ccgcatagc atcactgatg gcattcacag cctctatcac cagcccgctc 4140
accaccaac ataccctcct gtttaacatc ctggggggat gggtagccgc ccaacttgct 4200
cctccacagc ctgctctctg tttcgtaggc gccggcatcg ctggagcggc tgttggcagc 4260
ataggccttg ggaagtgct tgtggatatt ttggcagggt atggagcagg ggtggcaggc 4320
gcgctcgtgg cctttaaggc catgagcggc gagatgccct ccaccgagga cctggttaac 4380
ctactccctg ctatcctctc ccttggcgcc ctagtctgct gggctcgttg cgacgcgata 4440
ctgcgtcgcc acgtggggcc aggggagggg gctgtgcagt ggatgaaccg gctgatagcg 4500
ttcgtctcgc ggggtaacca cgtctccccc acgcaactat tgccctgagag cgacgctgca 4560
gcacgtgtca ctcatgctc ctctagtctt accatcactc agctgctgaa gaggcttcac 4620
cagtggatca acgaggactg gactgatttc tgctccggct cgtggctaag agatgttttg 4680
gattggatat gcacgggtgt gactgatttc ctctcatgt caacgtgggt acaaggaggt ctggcggggc 4740
cgattggcgg gattccctct ctctcatgt caacgtgggt acaaggaggt ctggcggggc 4800
gacggcatca tgcaaacacc ctgcccattgt ggagcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggcctgtggg gctcaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cccctgcacg cctctcccg gcgcaaatla ttctaggggc 4980
ctgtggcggg tggctgctga ggaatcctg gaggttacgc gggtagggga tttccactac 5040
gtgacgggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccgaaattc 5100
ttcacagaag tggatggggc gcggttgcac aggtacgctc cagcgtgcaa acccctccta 5160
cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
tgcgagcccg aaccggactg agcagtgtct acttccatgc tcaccgacc cccccacatt 5280
acggcggaga cggctaagcg taggtgccc aggggatctc cccctcctt ggccagctca 5340
tcagctagcc agctgtctgc gccctccttg aaggcaacat gcactaccgg tcatgactcc 5400
ccgacgctg acctcatcga ggccaacctc ctgtggcgcc agggatggg cggaacatc 5460
acccgctgg agtcagaaaa taaggtagta attttgact ctttcgagcc gctccaagcg 5520
gaggaggatg agagggaagt atccgttccg gcggagatcc tgcgagggtc caggaaattc 5580
cctcgagcga tgccccatg ggcacgccc gattacaacc ctccactgtt agagtccctg 5640
aaggaccgg actacgtccc tccagtggta caggggtgtc cattgccgcc tgccaaggcc 5700
cctccgatac cactccacg gagggaaggg acggttgtcc tgcagaaac taccgtgtct 5760
tctgccttgg cggagctcgc cacaagacc ttccggcagct ccgaatcgtc ggccgtcgac 5820
agcgccacgg caacggcctc tcctgaccag cctcccgagc acggcgacgc gggatccgac 5880
ggttgatcgt actctccat gcccccctt gagggggagc cgggggatcc cgatctcagc 5940
gacgggtctt ggtctaccgt aagcgaggag gctagtggg acgtcgtctg ctgctcgatg 6000
tcctacacat ggacagggcg cctgatcagc ccatgcgctg cgaggaaac caagtgcctc 6060
atcaatgcac tgagcaactc ttgtctcctg caccacaact tggctctatg tacaacatct 6120
cgacgcgcaa gctcgggca gaagaaggtc accttgaca gactgcaggt cctggacgac 6180
cactaccggg acgtgtcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
ctatccgttg aggaagcctg taagctgacg cccccacatt cggccagatc taaatttggc 6300
tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360
tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gagggtttct gcgtccaacc agagaagggg ggccgcaagc cagctcgctt tatcgtattc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
cctcaggccg tgatgggctc ttcatacggg ttccaatact ctccctggac gcgggtcgag 6600
ttctctggtg atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacacccgc 6660
tgttttgact caacggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg cccccgaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
gggggcccc tgactaatc taaagggcag aactgcggct atcgccggtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cggtaatacc ctcacatggt acttgaaggc cgctgcggcc 6900
tgtcgagctg cgaagctcca ggactgcacg atgctcgat cgggagacga ccttgtcgtt 6960
atctgtgaaa gcgcggggac ccaaggaggc gaggcgagcc tacgggcctt cacggaggct 7020
atgactagat actctgcccc ccctggggac ccgcccaaac cagaatacga cttggagtgt 7080
ataacatcat gctcctccaa tgtgtcagtc gcgcacgatg catctgcaa aagggtgtac 7140
tatctacccc gtgacccccc caccctcctt gcgcgggctg cgtgggagac agctagacac 7200
actccagtc atctcctggct aggcacatc atcatgtat cgccacctt gtgggcaagg 7260
atgatctga tgactcattt ctctctcacc cttctagctc aggaacaact tgaaaaagtc 7320
ctgattgttc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380
attcaacgac tccatggcct tagcgcatth tcaactcata gttactctcc aggtgagatc 7440
aatagggtg cttcatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500

```

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggtgc cacttgtggc 7560  
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620  
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680  
agcctgtctc gtgcccgacc ccgctgggtc atgtgggtgc tactcctact ttctgtaggg 7740  
gtaggcattc atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctggt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860  
ttttctectt ttttttctct ctttttttcc ttttctttcc tttgggtggc ccatcttagc 7920  
cctagtcacg gctagctgtg aaagggtccg gagccgcttg actgcagaga gtgctgatac 7980  
10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 4  
 <211> 7989  
 <212> DNA  
 <213> Hepatitis C Virus

5

<400> 4

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggg gactacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcca gtgcccgggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaagggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cgcccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggtctgt 480
ctgatgcgcg cgtgttccgg ctgtcagcgc agggcgcccc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacggcggt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcgg gggcaggatc tcctgtcacc tcacctgtgt cctgcggaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctaccctgc 780
cattcgacca ccaagcgaaa catcgcacg agcagcacg tactcggatg gaagccggtc 840
ttgtcgatca ggtgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttgc 900
ccaggtcaa ggcgcgatg cccgacggcg aggatctcgt cgtgacctat ggcgatgcct 960
gcttgccgaa tatcatgggt gaaatggccc gcttttcttg attcatcgac tgtggccggc 1020
tgggtgtggc ggaacgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080
ttggcgcgca atgggctgac cgcttctcgc tgctttacgg tatcgccgct cccgattcgc 1140
agcgcatcgc cttctatcgc cttcttgacg agttcttctg agtttaacaa gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgtc tggaataagg ccggtgtcgg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcgcgaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaagggaagc 1440
gttccctctg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500
aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg ccagaagggt accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gagggtaaaa 1740
aacgtctagg cccccgaac caccgagcag acgcgagggc tacttggctg catcatact 1860
atggcgctta ttacggccta ctcccaacag acgcgagggc tccaagtgtt ctccaccgca 1920
agcctcacag gccgggacag gaaccaggtc gagggggagg tcccaagtgtt ctccaccgca 1980
acacaatctt tcttgccgac ctgcgtcaat ggcgtgtgtt ggactgtcta tcatgtgtgc 2040
ggctcaaaga cccttgccgg cccaaagggg ccaatcacc aaatgtacac caatgtggac 2040
caggacctcg tcggctggca agcgcctccc gggcgcggtt ccttgacacc atgcacctgc 2100
ggcagctcgg acctttactt ggtcacgagg catgcccgat tcattccggt gcgcggcgcg 2160
ggcgacagca gggggagcct actctcccc agggccgtct cctacttgaa gggctcttgc 2220
ggcggtccac tgctctgccc ctcggggcac gctgtgggca tcttctgggc tgccgtgtgc 2280
acccgagggg ttgcgaaggc ggtggacttt gtaccgctcg agtctatgga aaccactatg 2340
cggtccccgg tcttcacgga caactcgtcc cctccggcgg taccgcagac attccagggt 2400
gcccattctc acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460
gcccagggtg ataaggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 2520
gcgtatatgt ctaaggcaca tggatcgac cctaactca gaaccggggg aaggaccatc 2580
accacgggtg cccccatcac gtactccacc tatggcaagt ttcttgccga cgggtgttgc 2640
tctggggcg cctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 2700
atcctgggca tcggcacagt cctggacca ggcgagacgg ctggagcgcg actcgtcgtg 2760
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggagggt 2820
gctctgtcca gcactggaga aatccccctt tatggcaag ccatccccat cgagaccatc 2880
aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccggc 2940
aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc 3000
ataccaacta gcggagacgt cattgtcgta gcaacggagc ctctaatagc gggctttacc 3060
ggcgatttcg actcagtgat cgactgcaat acatgtgtca ccagacagt cgacttcagc 3120
ctggaccgca ccttcacatc tgagacgacg accgtgccac aagacgcggg gtcaacgtcg 3180
cagcgcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga 3240
gaacggccct cgggcatgtt cgattcctcg gttctgtgct agtgcctatga cgcgggctgt 3300
gcttggtagc agctcacgcc ccgcgagacc tcagttaggt tgccgggctta cctaaacaca 3360
ccagggttgc ccgtctgcca ggaccatctg gatttctggg agagcgtctt tacaggcttc 3420
accacatag acgcccattt cttgtccacg actaagcagg caggagacaa ctccctctac 3480
ctggtagcat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac 3540

```

caaatgtgga agtgtctcat acggctaaag cctacgtctg acggggccaac gcccctgctg 3600  
 tataggctgg gagccgttca aaacgaggtt actaccacac acccataaac caaatacatc 3660  
 atggcatgca tgtcgggtga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3720  
 gtccctagcag ctctggccgc gtattgcctg acaacaggca gcgtgggtcat tgtgggcagg 3780  
 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccgggagttc 3840  
 gatgagatgg aagagtgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc 3900  
 gaacaattca aacagaaggg aatcgggttg ctgcaaacag ccaccaagca agcggaggct 3960  
 gctgctcccg tgggtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg 4020  
 tggaaattca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080  
 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgtctac caccacaatc 4140  
 accctcctgt ttaacatcct ggggggatgg gtggccgccc aacttgctcc tcccagcgct 4200  
 gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260  
 aaggtgcttg tggatatttt ggaggttat ggagcagggg ttggcagggc gctcgtggcc 4320  
 ttttaaggta tgagcggcga gatgcctccc accgaggacc ttggttaacct actccctgct 4380  
 atcctctccc ctggcgcctc agtcgtcggg gtcgtgtgct cagcgatact gcgtcggcac 4440  
 gtggggccag gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcggcg 4500  
 ggttaaccag tctccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 4560  
 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620  
 gaggactgct ccacgcatg ctccggctcg ttgctaagag atgtttggga ttggatatgc 4680  
 acggtgttga ctgatttcaa gacctggctc cagtccaagc tcttgcgcgc attgcccggg 4740  
 gtccccctct tctcatgtca acgtgggtac aagggagtct ggccggggcg cgcatcatg 4800  
 caaaccacct gccatgtgg agcacagatc accggacatg tgaanaacgg ttccatgagg 4860  
 atcgtggggc ctaggacctg tagtaaacag ttgcatggaa cattccccat taacgcgtac 4920  
 accacgggcc cctgcacgcc ctccccggcg ccaaatattt ctaggcgctg gtggcgggtg 4980  
 gctgctgagg agtacgtgga ggttacgcgg gtgggggatt tccactacgt gacgggcatg 5040  
 accactgaca acgtaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg 5100  
 gatgggggtg ggttgacag gtacgtccca gcgtgcaaac cctcctacg ggaggaggtc 5160  
 acattcctgt tcgggtcaca tcaatacttg gttgggtcac agtccccatg cgagcccgaa 5220  
 ccggacgtag cagtgtctac ttccatgtct accgacccct cccacattac ggcgagacg 5280  
 gctaagcgta ggctggccag gggatctccc cctccttgg ccagctcacc agctagccag 5340  
 ctgtctgcgc ctcccttga actaccgctc atgactcccc gacgctgac 5400  
 ctcatcgagg ccaacctct gtggcggcag gagatggcg ggaacatcac ccgctggag 5460  
 tcagaaaaata aggtagtaat tttggactct ttccagccgc tccaagcgga ggaggatgag 5520  
 agggaaagtat ccgttccggc ggagatcctg cggaggtcca ggaattcccc tcgagcgatg 5580  
 cccatattgg cacgcccga ttacaacct ccaactgttag agtccctgaa ggacccggac 5640  
 tacgtccctc cagtgtgaca cgggtgtcca ttgccgctg ccaaggcccc tccgatacca 5700  
 ctccacgga ggaagaggac ggtgtctctg tcagaatcta ccgtgtcttc tgccttggcg 5760  
 gagctcgcca caaagacctc cggcagctcc gaatcgtcgg ccgtcgacag cggcacggca 5820  
 acggcctctc ctgaccagcc ctccgacgac ggccgacggg gatccgacgt tgagtcgtac 5880  
 tctccatgc ccccccttga gggggagccg ggggatcccc atctcagcga cgggtcttgg 5940  
 tctaccgtaa gcgaggaggc tagtgaggac gtcgtctgct gctcgatgct ctacacatgg 6000  
 acaggcgccc tgatcacgcc atgcgtcgcg gaggaacca agctgcccat caatgcactg 6060  
 agcaactctt tgctccgtca ccacaacttg gtctatgcta caacatctcg cagcgcaagc 6120  
 ctgcggcaga agaaggtcac ctgtgacaga ctgcaggtcc ttgacgacca ctaccgggac 6180  
 gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaaactct atccgtggag 6240  
 gaagcctgta agctgacgcc cccacattcg gccagatcta aatttggcta tggggcaaa 6300  
 gacgtccgga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg 6360  
 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 6420  
 gtccaaccag agaagggggg ccgcaagcca gctcgcttca tcttattccc agatttgggg 6480  
 gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccacctccc tcaggccgtg 6540  
 atgggctctt catacggatt ccaatactct cctggacagc ggttcgagtt cctggtgaat 6600  
 gcctggaaag cgaagaaatg ccctatgggc ttgcgatatg acaccgctg ttttgactca 6660  
 acggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc 6720  
 cccgaagcca gacaggccat aaggtcgtc acagagcgcc ttacatcgg gggccccctg 6780  
 actaattcta aagggcagaa ctgcggctat cgccggtgcc gcgcgagcgg tgtactgacg 6840  
 accagctcgg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgcg 6900  
 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 6960  
 gcgggggacc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac 7020  
 tctgcccccc ctggggaccc gcccaacca gaatacgact ttgagttgat aacatcatgc 7080  
 tcttccaatg tgcagtcgc gcacgatgca tctggcaaaa ggtgtacta tctcaccctg 7140  
 gacccccacca ccccccttgc gcgggctgcg tgggagacag ctagacacac tccagtcaat 7200  
 tcttggttag gcaacatcat catgtatgcg cccaccttgt gggcaaggat gatcctgatg 7260  
 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag 7320  
 atctacgggg cctgttactc cttgagacca cttgacctac ctgagatcat tcaacgactc 7380  
 catggcctta gcgcattttc actccatagt tactctccag gtgagatcaa taggggtggc 7440  
 tcatgcctca ggaaacttgg ggtaccgccc ttgcgagctc ggagacatcg ggccagaagt 7500

gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 7560  
aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 7620  
5 ttatccagct ggttcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt 7680  
gcccgaaccc gctggttcat gtggtgccta ctccacttt ctgtaggggt aggcattctat 7740  
ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgtttt 7800  
tttccctttt ttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt 7860  
tttttcctct ttttttcctt ttctttcctt tgggtgctcc atcttagccc tagtcacggc 7920  
tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 7980  
10 agatcaagt 7989

15

20

25

30

35

40

45

50

55

<210> 5  
 <211> 8649  
 <212> DNA  
 <213> Hepatitis C Virus

5

&lt;400&gt; 5

10

15

20

25

30

35

40

45

50

```

gccagccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtgggtctg cggaaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgoga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aaccaaacg ggcgcgcat gattgaacaa gatggattgc 420
acgcagggttc tcgcgcgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgcgtgtgc actgaagcgg 660
tggaaggactg gctgctattg ggcaagtgcc cggggcagga tctcctgtca tctcaccttg 720
ctcctgcoga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctactcg cccattcgac caccagcgca aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcgcgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccc aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggtg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcatc gccttctatc gccttcttga caggttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgtttatt 1320
tccaccatat tgcgtctttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtgtc ttccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaa cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcgaccttc 1500
tgcaaggcag ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaagg ggcacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgccagaag 1680
gtaccccat ttatgggata tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggaccg ggagatggca gcatcgtgcg gaggcggcgt tttcgtaggt 1860
ctgatactct tgacctgtc accgcactat aagctgttcc tcgctaggct catatggttg 1920
ttacaatat ttatcaccag ggccgaggca cacttgcaag tgtggatccc cccctcaac 1980
gttcgggggg gccgcgatgc cgtcatctc ctacgtgcg cgtatccacc agagctaata 2040
tttaccatca ccaaaatctt gctcgccata ctcggtccac tcatgggtct ccaggctggt 2100
ataaccaaa tgccgtactt cgtgcgcgca cacgggctca ttctgtcatg catgctggtg 2160
cggaagggtg ctgggggtca ttatgtccaa atggctctca tgaagttggc cgcactgaca 2220
ggtacgtacg tttatgacca tctacccca ctgcgggact gggccacgc gggcctacga 2280
gaccttgccg tggcagttga gcccgctgct ttctctgata tggagaccaa ggttatcacc 2340
tggggggcag acaccgcggc gtgtggggac atcatcttgg gcctgcccgt ctccgccgcg 2400
agggggaggg agatacatct gggaccggca gacagccttg aaggcgagg gtggcgactc 2460
ctcgcgcta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 2520
agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtggc ctccaccgca 2580
acacaatctt tcctggcgac ctgcgtcaat ggctgtgtt ggactgtcta tcatggtgcc 2640
ggctcaaaga cccttgccg cccaagggc ccaatcacc aaatgtacac caatgtggac 2700
caggacctcg tcggctggca agcgcccccc ggggcgcgtt ccttgacacc atgcacctg 2760
ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggt gcgcggcgg 2820
ggcgacagca gggggagcct actctcccc aggcgcgtct cctacttgaa gggctcttcg 2880
ggcggtccac tgctctgccc ctgggggac gctgtgggca tctttcgggc tgccgtgtgc 2940
acccgagggg ttgcgaaggc ggtggacttt gtaccgctcg agtctatgga aaccactatg 3000
cggctccccg tcttcacgga caactcgtcc cctccggccg tacccgagac attccagggtg 3060
gcccatctac acgcccctac tggtagcggc aagagcacta aggtgccggc tcggtatgca 3120
gcccagggtg ataagtgct tgcctgaac ccgtccgtcg ccgccacct aggtttcggg 3180
gcgtatatgt ctaaggcaca tggatcgac cctaacatca gaaccgggt aaggaccatc 3240
accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtgtgtg 3300
tctggggcgg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgacctc 3360
atcctgggca tcggcacagt cctggaccaa cgtggagcgg ctggagcgcg actcgtcgtg 3420
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 3480
gctctgtcca gcactggaga aatccccctt tatggcaaa ccatcccat cgagaccatc 3540

```

55

aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgag 3600  
 aagctgtccg gcctcggact caatgtctgta gcatattacc ggggcccctga tgtatccgtc 3660  
 ataccaacta gcggagacgt cattgtcgtta gcaacggacg ctctaagac gggctttacc 3720  
 5 ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc 3780  
 ctggaccoga ccttcacat tagagacgac accgtgccac aagacgcggt gtcacgctcg 3840  
 cagcggcgag gcaggactgg taggggcagg atgggcattt acagggttgt gactccagga 3900  
 gaacggccct cgggcatgtt cgattcctcg gttctgtgag agtgctatga cgcgggctgt 3960  
 gcttggtacg agctcacgcc cgccgagacc tcagttaggt tgccggctta cctaaacaca 4020  
 ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggccctc 4080  
 10 accccatag acgccattt cttgtcccag actaagcagg caggagacaa cttccctac 4140  
 ctggtagcat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac 4200  
 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaaac gccctctgtg 4260  
 tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc 4320  
 atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 4380  
 gtccatagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 4440  
 atcatcttgt ccggaagacc ggccatcatt cccgacaggg aagtccttta ccgggagttc 4500  
 gatgagatgg aagatgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc 4560  
 15 gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccagca agcggaggct 4620  
 gctgctcccg ttgtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg 4680  
 tggaaattca tcagcgggat acaatattta gcaggctgtt gcaactcgcc tggcaacccc 4740  
 gcgatagcat cactgatggc attcacagcc tctatacca gcccgctcac caccacaat 4800  
 accctcttgt ttaacatcct ggggggatgg gtggccgccc aacttgctcc tccagcgct 4860  
 gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4920  
 20 aaggtgtctg ttgatatattt ggacaggttat ggagcagggg ttggcagcgc gctcgtggcc 4980  
 tttaaagtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 5040  
 atcctctccc cttggccctt agtcgtcggg gtcgtgtgag cagcgatact gcgtcggac 5100  
 gtgggcccag gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgagg 5160  
 ggtaaccacg tctcccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 5220  
 cagatctctt ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 5280  
 25 gaggactgct ccacgccatg tcccgctcg tggtctaaag atgtttggga ttggatatgc 5340  
 acgggtgtga ctgatttcaa gacctggctc cagtcgaagc tctgcccgcg attgcccggg 5400  
 gtccctctct tctcatgtca acgtgggtac aagggagctt ggccggggcga cggcatcatg 5460  
 caaacacct gccatgtgg agcacagatc accggacatg tgaanaacgg ttccatgagg 5520  
 atcgtggggc ctaggacctg tagtaacacg ttggcatggaa cattccccat taacgcgtac 5580  
 accacgggcc cctgcacgcc ctccccggcg ccaaatattt ctaggcgctg gtggcgggtg 5640  
 gctgctgagg agtacgtgga ggttacgcgg gtgggggatt tccactacgt gacgggcatg 5700  
 30 accatgaca acgtaaagt cccgtgtcag gttccggccc ccgaattctt cacagaagtg 5760  
 gatgggggtg ggttgacacg gtacgctcca gcgtgcaaac ccctcctacg ggaggaggtc 5820  
 acattcctgg tcgggctcaa tcaatacctg gttgggtcac agctcccatg cgagcccga 5880  
 ccggacgtag cagtgctcac ttccatctc accgacccct cccacattac ggccgagacg 5940  
 gctaagcgta ggttgccag gggatctccc ccctccttgg ccagctcacc agctagccag 6000  
 ctgtctcgcc ctcccttgaa ggcaacatgc actacccgtc atgactcccc ggacgctgac 6060  
 35 ctcatcgagg ccaacctcct gtggcgagc gagatggcg ggaacatcac ccgctggag 6120  
 tcagaaaata aggtagtaat tttggactct ttcgagccgc tccaagcggg ggaggatgag 6180  
 agggaaagtat ccgttccggc ggagatcctg cggaggtcca ggaaattccc tcgagcgtatg 6240  
 cccatatggg cacgcccggg ttacaacctt ccactgttag agtccggaa ggacccggac 6300  
 tacgtccctc cagtggatga cgggtgtcca ttgccgctg ccaaggcccc tccgatacca 6360  
 cctccacgga ggaagaggac ggttgtctcg tcagaatcta ccgtgtcttc tgccttggcg 6420  
 gagctcgcca caaagacctt cggcagctcc gaatcgtcgg ccgtcgacag cggcacggca 6480  
 40 acggccctct ctgaccagcc ctccgacgag ggccgacgag gatccgacgt tgagtgtac 6540  
 tctccatgct ccccccttga gggggagccg ggggatcccg atctcagcga cgggtcttgg 6600  
 tctaccgtaa gcgaggaggc tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg 6660  
 acaggcgccc tagtcacgcc atgcgctgag gaggaacca agctgcccat caatgcactg 6720  
 agcaactctt tgctccgtca ccacaacttg gtctatgcta caacatctcg cagcgcaagc 6780  
 ctgcggcaga agaaggtcac ctttgacaga ctgcaggctc ttgacgacca ctaccgggac 6840  
 45 gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaaactct atccgtggag 6900  
 gaagcctgtg agctgacgcc cccacattcg gccagatcta aatttggcta tggggcaaa 6960  
 gacgtccgga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg 7020  
 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 7080  
 gtccaaccag agaagggggg ccgcaagcca gctcgcccta tctgattccc agatttgggg 7140  
 gttcgtgtgt gcgagaaaat gggcccttac gatgtggtct ccacctccc tcaggccgtg 7200  
 atgggctctt catacggatt ccaataactct cctggacagc gggctcaggtt cctggtgaat 7260  
 50 gcttggaagg cgaagaatg cctatgggc ttccgatatg acaccgctg ttttgactca 7320  
 acggtcactg agaatgacat ccgtgtgag gactcaatct accaatgttg tgacttggcc 7380  
 cccgaagcca gacaggccat aaggtcgtc acagagcgcc tttacatcgg gggcccccctg 7440  
 actaatctta aagggcagaa ctgcggctat cgccgggtgc gcgcgagcgg tgtactgacg 7500

accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgcg 7560  
 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 7620  
 gcggggaccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac 7680  
 5 tctgcccccc ctgggggaccc gcccaaacca gaatacgact tggagttgat aacatcatgc 7740  
 tcttccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgtacta tctcaccctg 7800  
 gacccccacca ccccccttgc gcgggctgcg tgggagacag ctagacacac tccagtcaat 7860  
 tcctggctag gcaacatcat catgtatgcg cccaccttgt gggcaaggat gatcctgatg 7920  
 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag 7980  
 10 atctacgggg cctgttactc cattgagcca cttgacctac ctcagatcat tcaacgactc 8040  
 catggcctta gcgcattttc actccatagt tactctccag gtgagatcaa taggggtggct 8100  
 tcatgcctca ggaaacttgg ggtaccgccc ttgcgagtct ggagacatcg ggccagaagt 8160  
 gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 8220  
 aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 8280  
 ttatccagct gggttcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt 8340  
 15 gcccgaaccc gctggttcat gtggtgccta ctctacttt ctgtaggggt aggcattctat 8400  
 ctactcccca accgatgaac ggggagctaa aactccagg ccaataggcc atcctgtttt 8460  
 tttccctttt ttttttctt ttttttttt ttttttttt ttttttttt ttctcctttt 8520  
 tttttcctct ttttttcctt ttctttcctt tgggtgctcc atcttagccc tagtcacggc 8580  
 tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 8640  
 20 agatcaagt 8649

25

30

35

40

45

50

55

<210> 6  
 <211> 8001  
 <212> DNA  
 <213> Hepatitis C Virus

5

<400> 6

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctcgtcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tccctttctg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttcgca gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgcgca gaaagtatcc atcatggctg atgcaatgcg gcgctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaaagcgg tcttctgatc caggatgacg tggacgaaga gcatcagggg ctgcgcagcg 900
ccgaactgtt cgccaggctc aagcgcgca tgcgcgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg cgtgggtgtg gcggaccgct atcaggacat agcgttggct acccgatgata 1080
ttgctgaaga gcttgccggc gaattgggtg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagttttaa 1200
cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagcgg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tcaccatat tgcctgtctt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tccccctctc gccaaaggaa tgcaaggtct gttgaatgta 1440
gtgaaggaa cagttctctt ggaagctctt tgaagacaaa caacgtctgt agcgaccctt 1500
tgagggaagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaggcg ggcacaaccc cagtgcacag ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccattc gtatgggacg tgcctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta gggcccccg accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc agacgcgagg cctactttgc 1860
tgcatcatca ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagt 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcggtg ttggactgtc 1980
tatcatggtg ccggctcaaa gacccttgcc ggcccaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggtgg caagcgcccc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
gtgcgcggcg ggggcagacg cagggggagc ctactctccc ccaggcccgt ctcctacttg 2220
aagggtctct cggcggttcc actgctctgc cctcggggc atgctgtggg catctttcgg 2280
gctgccgtgt gcaaccgagg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460
gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgctcgt gcgcgccacc 2520
ctagggttctg gggcgtatat gtctaaggca catggtatcg accctaacat cagaaccggg 2580
gtagggacca tcaccacggg tgcctccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgctatgac atcataatat gtgatgagtg ccactcaact 2700
gactcgacca ctatctggg catcggcaca gtctggacc aagcggagac ggctggagcg 2760
cgactcgctg tgcctgccac cgctacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcactgga gaaatcccc tttatggcaa agccatcccc 2880
atcgcgacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagctcgccg cgaagctatc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcgagac gtcattgtcg tagcaacgga cgtctaatg 3060
acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggaccc gaccttacc attgagacga cgaccgtgcc acaagacgag 3180
gtgtcacgct cgcagcgcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactcag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgtat 3300
gacgcggtct gtgcttggta cgcgtcacg ccgcgcgaga cctcagttag gttgcgggct 3360
tacctaaaca caccagggct gcccgctctc caggaccatc tggagttctg ggagagcgct 3420
tttacaggcc tcaccacat agacgcccct ttcttctccc agactaggca ggcaggagac 3480
aacttcccc acctggtagc ataccaggct acggtgtgcg ccagggtcga ggctccacct 3540

```

	ccatcggtgg	acccaaatgtg	ggagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgccccctgc	tgtataggct	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcgggt	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtagggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
	attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccgacag	ggaagtcttt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgtgtctcc	ctgtgtggaa	tccaagtggc	ggaccctcga	agccttcttg	4020
	gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcagcgctt	gtccactctg	4080
	cctggcaaac	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgctc	4140
10	accaccaaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctcccagcg	ctgcttctgc	tttcgtagge	gccggcatcg	ctggagcggc	tggtggcagc	4260
	ataggcccttg	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
	gcgctcgctg	cctttaagggt	catgagcggc	gagatgccct	ccaccgagga	cctgggtaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctgcgtcggc	acgtggggcc	agggggaggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
	ttcgcttcgc	ggggtaacca	cgctccccc	acgcactatg	tgctgagag	cgcagctgca	4560
15	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggtctcac	4620
	cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgttttg	4680
	gattggatga	gcacggtggt	gactgatttc	aagacctggc	tccagtccaa	gctcctggcg	4740
	cgattcccg	gagtcctctt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaacacc	ctgcccatgt	ggagcacaga	tcaccggaca	tgtagaaaaa	4860
	ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
20	atcaacgcgt	acaccacggg	ccctcgacg	ccctccccgg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtgtctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggga	tgaccactga	caacgtaaag	tgcccggtgc	aggttccggc	ccccgaattc	5100
	ttcacagaa	tgatgggggt	gcggttgca	aggtacgctc	cagcgtgcaa	accctctcac	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
	tgccggcccg	aaccggacgt	agcagtgctc	acttccatgc	tcaccgaccc	ctccacattt	5280
25	acggcgggga	cggttaagcg	taggctggcc	aggggatctc	ccccctcctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccctg	aaggcaacat	gcactaccgg	tcagtactcc	5400
	ccggacgctg	acctcatcga	ggccaaacct	ctgtggcggc	aggagatggg	cgggaacatc	5460
	acccgcgtgg	agtcagaaaa	taaggtagta	attttggtg	ctttcgagcc	gctccaagcg	5520
	gagggagatg	aggaggaaag	atccgttccg	gcggagatcc	tgccggaggc	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgcccc	gattacaacc	ctccactggt	agagtctctg	5640
	aaggacccgg	actacgtccc	tccagtggta	cacgggtgtc	cattgcggcc	tgccaaggcc	5700
30	cctccgagcc	cacctccacg	gaaggagagg	acggttgctc	tgtagaate	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgct	ggccgtcgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	ggtgagtcgt	actcctecat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gctagtggag	acgtcgtctg	ctgctcgatg	6000
	tcctacacat	ggacaggcgc	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
	gtcaatgcac	tgagcaactc	tttgctccgt	caccacaact	tggtctatgc	tacaacatct	6120
35	cgcagcgcaa	gcctgcggca	gaagaagggt	acctttgaca	gactgcagg	cctggagcac	6180
	cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacattt	cggccagatc	taaatattgg	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgtccgtg	6360
	tggaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaacc	agagaagggt	ggccgcaagc	cagctcgctt	tatcgtattc	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
	cctcaggccc	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctggtga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tggtttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtagtcttg	ccccgaagc	cagacaggcc	ataaggctcg	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaatcc	taaagggcag	aactgcggct	atcgccgggt	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggtaatacc	ctcacatggt	atttgaagcg	cgtgcggccc	6900
	tgctcagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgcgtgt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gaggcgagcc	tacgggcctt	cacggagcgt	7020
	atgactagat	actctgcccc	ccctggggac	ccgccccaac	cagaatacga	cttgaggttg	7080
	ataacatcat	gctcctccaa	tggtgcagtc	gcgcacgatg	catctggcaa	aaggggtgac	7140
	tatctcaccc	gtgacccacc	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
	actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgccacactt	gtgggcaagg	7260
50	atgactcctga	tgactcatct	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagatgtgc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560  
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620  
tcccagttgg atttatccag ctgggttcgtt gctggttaca gcgggggaga catatatcac 7680  
agcctgtctc gtgcccgacc ccgctgggtc atgtgggtgcc tactcctact ttctgtaggg 7740  
gtaggcattc atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860  
ttttctcctt tttttttcct ctttttttcc ttttcttcc tttgggtggc ccatcttagc 7920  
cctagtcaag gctagtgtg aaagggtccg gagccgcttg actgcagaga gtgctgatac 7980  
10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 7  
 <211> 11076  
 <212> DNA  
 <213> Hepatitis C Virus

5

&lt;400&gt; 7

10

15

20

25

30

35

40

45

50

```

gccagccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc ggagagacca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgacgggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcggtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gacogtgcac catgagcacg aatcctaaac 360
ctcaagaaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggctc tcgggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgc actgaagcgg 660
gaaaggactg gctgctattg ggcaagtgcc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctactcg cccattcgac caccagcgca aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgcca 900
ccgaactgtt cgccaggctc aaggcgcgca tggccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200
cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttaact gcgcaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgcgctctt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt ttccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaa gcttctctc ggaagcctct tgaagacaaa caacgtctgt agcgacctt 1500
tgagggcagc ggaacccccc acctggcgac aggtgcctct gcggcctaaa gccacgtgta 1560
taagatacac ctgcaaaagg ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgccagaag 1680
gtaccccatg gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaaagctta ggcaccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatccataa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttccggggc ggtggctaga tgcgtgggtg agtttaacct 1920
ttgcgcgca gggggcccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccg ggtacccctg gccctctat ggcaatgagg gcttggggtg ggcaggatgg 2100
ctcctgtcac ccggtggctc tcggcctagt tggggcccca cggaaccccg gcgtaggctg 2160
cgcaatttgg gtaaggctcat cgataccctc acgtgcggct tcgccgatct catggggtac 2220
attccgctcg tgcgcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280
gttctggagg acggcggtga ctatgcaaca gggaatctgc ccggttgctc ctttctatc 2340
ttcttttgg ctttctgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cagcaacgac tgcctcaacg caagcattgt gtatgaggca 2460
gcggaatga tcatgcatac ccccggtgct gtgcctcgcc ttcgggagaa caactcctcc 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt gggggcgctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgcccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacagggtac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgcgtgggat atggtggcgg gggcccatg gggagtccta 2880
gcgggccttg cctactatc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaaacc 3000
ctcgggatta cgtccctctt ttaccgggg tcattcccaga aatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgcc agagcgcatg 3180
gccagctgca gcccacatga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccgggccc gtgcggtatc 3300
gtaccgcgcy cgcaggtgtg tggctcagtg tactgcttca ccccaagccc tgcgtgggtg 3360
gggacgacgc accggttcgg cgtccctacg tacagttggg gggagaatga gacggagctg 3420
ctgcttctta acaacacgcy gcccgccgaa ggcaactggg ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccggtgta acatcggggg gatcggaat 3540

```

55

5

10

15

20

25

30

35

40

45

50

55

```

aaaacacctga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
tgtgtgttcgg ggccttgggt gacacccaga tgcttgggtc actaccata caggcttttg 3660
cactaccocct gcactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tcgaagccgc atgcaattgg actcaggag agcgttga cctggaggac 3780
agggacagat cagagcttag ccgcgtgctg ctgtctacaa cggagtggca ggtattgccc 3840
tgttccttca ccaccctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
gtggacgtac aatacctgta cggatagagg tccggcgttg tctcctttgc aatcaaatgg 3960
gagtattgcc tgttgctctt ccttcttctg gcggacgcgc gcgtctgtgc ctgcttgttg 4020
atgatgctgc tgaatgctca agctgaggcc gccctagaga acctggtggt cctcaacgcg 4080
gcatccgtgg ccggggcgca tggcattctc tccttctctg tgttctctg tctgctctgg 4140
tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacgcgt atggccgcta 4200
ctctgctccc tctggtggtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
tcgtgctggg gcgcgttttt cgtaggctctg atactcttga ccttgtcacc gcactataag 4320
ctgttctctg ctaggctcat atggtgggta caatatttta tcaccagggc cgaggcacac 4380
ttgcaagtgt ggtatccccc gctaatcttt accatcacca aaatcttgc ctccctctc 4440
acgtgctgca tcacaccaga gctaatcttt accatcacca aaatcttgc ctccctactc 4500
gggtccactca tgggtctcca ggtgtgtata accaaagtgc cgtacttctg gcgcgcacac 4560
gggtctcattc gtgcatgcat atggtgtcgg aaggttgcgt ggggtcatta tgtccaaatg 4620
gctctcatga agttggccgc actgacaggt acgtacgttt atgacctct caccacctg 4680
ggggactggg ccacgcgggg cctacgagac ctgtcgttgg cagttgagcc cgtcgtcttc 4740
tctgatattg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atcttggggc tgcctgtctc cgcccgaggg gggaggagga tacatctggg accgagac 4860
agccttgaag ggcagggggt gcgactcttc gcgctatta cggcctactc ccaacagacg 4920
cgaggctctc ttggctgcat catcactagc ctcacaggcc gggacaggaa ccaggctcag 4980
ggggagggtc aagtgtctc caccgcaaca caatcttcc tggcgacctg cgtcaatggc 5040
gtgtgttga ctgtctatca tgggtccggc tcaaagacc ttgcccggcc aaagggccca 5100
atcacccaaa gtacacca caactcgtcg gacctcgtc gctggcaagc gcccccggt 5160
gcgcgttctc tgacacctg cacctgcggc agctcggacc tttacttggg caccaggcat 5220
gcatgtgca ttccggtgcg ccgctctc gacagcagg ggagcctact ctccccagg 5280
ccgctctctc acttgaaggg ctctcgggc ggtccactgc tctgcccctc ggggcatgct 5340
gtggcctatc ttccggctgc cgtgtgcacc cgagggttgc cgaaggcggt ggactttgta 5400
cccgtcgagt ctatggaac cactatgcgg tccccggtc tcacggacaa ctccgtccct 5460
ccggccgtac cgacagacct ccagggtggc catctacacg ccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gtatgcagcc caagggtata aggtgcttg cctgaacccg 5580
tccgtcgccg ccaccctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640
aacatcagaa ccggggtagg gacctacacc acgggtgccc ccatcacgta ctccacctat 5700
ggcaagtctc ttgccgacgg gggggcgccct atgacatcat aatatgtgat 5760
gagtgcact caactgactc gacctatc ctgggcatcg gcacagtcct ggaccaagcg 5820
gagacggctg gagcgcgact cgtctgtctc gccaccgcta cgcctccggg atcggtcacc 5880
gtgcccactc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaagacca tccccatgc gacctcaag ggggggaggg acctcatttt ctgccattcc 6000
aagaagaaat gtgatgagct atccgtcata gacctcagc cagactcaa tgcgttagca 6060
tattaccggg gctttagtgt cttacccggc gatttcgact cagtgatcga ctgcaatata 6120
acggacgctc taatgacggg cttacccggc gacctcagc cagtgatcga ctgcaatata 6180
tgtgtcacc agacagtcga cttacccggc gacctcagc cagtgatcga ctgcaatata 6240
gtgccacaag acgcgggtgc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcca ttctcggtt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca gggctgccc tctgcccagg ccatctggag 6480
ttctgggaga gctctttac aggtctcacc cacatagacg cccatttctt gtcccagact 6540
aggcaggcag gagacaactt cccctacctg gtagcatacc aggtacggg gtgcgccagg 6600
gctcaggctc cacctccatc gtgggaccaa atgtgggagt gtctcatacg gctaaagcct 6660
acgctgcacg ggccaacgcc cctgctgtat aggtcgggag cgttcaaaa cgaggttact 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
acgagcacct ggtgctggtt aggcggagtc ctacgagctc tggcccgcta ttgctgaca 6840
acaggcagcg tggctattgt gggcaggatc atcttgtccg gaaagccggc catcattccc 6900
gacagggaag tcctttaccg ggagttcgat gagatggaag agtgccctc acacctccct 6960
tacatcgaac agggaaatga gctcggcgaa caattcaaac agaaggcaat cgggttgctg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaaatcaa gtggcgagc 7080
ctcgaagcct tctggcgaa gcatatgtgg aatttcatca gcgggatata atatttagca 7140
ggcttgtcca ctctgctgg caaccccgcg atagcatcac tctctgttta acatcctggg gggatgggtg 7200
atcacacagc cgctcaccac ccaacatacc ctcctgttta tctgctttgc taggcgcggc catcgctgga 7260
gccgcccaac ttgctctccc cagcgtgct tctgctttgc taggcgcggc catcgctgga 7320
gcggctgttg gcagcatagg ccttgggaag gtgcttgggt atattttggc aggttatgga 7380
gcagggggtg caggcgcgct cgtggccttt aaggtcatga gcggcgagat gccctccacc 7440
gaggacctgg ttaacctact ccctgctatc ctctccctg gcgccttagt cgtcggggtc 7500

```

5 gtgtgcgcag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560  
 aaccggctga tagcgttcgc ttcgcggggg aaccacgtct ccccccacgca ctatgtgtct 7620  
 gagagcgacg ctgcagcagc tgtcactcag atcctctcta gtcttaccat cactcagctg 7680  
 ctgaagaggc ttcaccagtg gatcaacgag gactgtctca cgccatgctc cggctcgtgg 7740  
 ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800  
 tccaagctcc tgcgcgatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860  
 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggagc acagatcacc 7920  
 ggacatgtga aaaaacgggtc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980  
 catggaacat tcccattaa cgcgtacacc acgggcccct gcacgccctc cccggcgcca 8040  
 10 aattattcta gggcgctgtg gccgggtggct gctgaggagt acgtggagggt tacgcgggtg 8100  
 ggggatttcc actacgtgac gggcatgacc actgacaacg taaagtgccg gtgtcaggtt 8160  
 ccggcccccg aattcttcac agaagtggat ggggtgcggg tgcaacaggtg cgctccagcg 8220  
 tgcaaaacccc tcctacggga ggaggtcaca ttcttggtcg ggctcaatca atacctggtt 8280  
 ggggtcacagg tcccatgcgg gcccgaaacc gacgtagcag tgcctacttc catgtctacc 8340  
 gacccctccc acattacggc ggagacggct aagcgtaggc tggccagggg atctcccccc 8400  
 15 tccttggcca gctcatcagc tagccagctg tctgcgcctt ccttgaaggc aacatgcact 8460  
 acccgtcatg actccccgga cgtgacctc atcgaggcca acctcctgtg gcggcaggag 8520  
 atggcgcgga acatcacccc cgtggagtcg gaaaataagg tagtaatttt ggactcttcc 8580  
 gagccgctcc aagcggagga ggtgagagg gaagtatccg ttccggcgga gatcctgcgg 8640  
 aggtccagga aattccctcg agcgtgccc atatgggcac gcccggatta caaccctcca 8700  
 ctgttagagt cctggaagga cccggactac gtccctccag tggtaacagc gtgtccattg 8760  
 20 ccgttcagga aggcctccc gatccacct ccacggaggg agaggacggt tgtcctgtca 8820  
 gaatctaccg tgtcttctgc cttggcgagg ctcgccacaa agaccttcgg cagctccgaa 8880  
 tcgtcggccg tcgacagcgg cacggcaacg gcctctctcg accagccctc cgacgacggc 8940  
 gacgcgggat ccgacgttga gtcgtactcc tccatgcccc cccttgaggg ggagccgggg 9000  
 gatcccgatc tcagcgacgg gtcttggtct accgtaagcg aggaggctag tgaggacgctc 9060  
 gtcgtctgct cgatgtccta cacatggaca ggcgcctga tcacgccatg cgctgcggag 9120  
 25 gaaaccaagc tgcccgtaa tgcactgagc aactctttgc tccgtacca caacttggtc 9180  
 tatgtacaaa catctcgag cgcaagcctg cggcagaaga aggtcacctt tgacagactg 9240  
 caggctcctg acgaccacta cccggagcgt ctcaaggaga tgaaggcga ggcgctccaca 9300  
 gttaaggcta aacttctatc cgtggaggaa gcctgtaagc tgacgcccc acattcggcc 9360  
 agatctaaat ttggctatgg ggcaaaggac gtccggaacc tatccagcaa ggccgttaac 9420  
 cacatccgct ccgtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480  
 30 atcatggcaa aaaaagaggt ttctgcgtc caaccagaga aggggggccc caagccagct 9540  
 cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaaatggc cctttacgat 9600  
 gtggtctcca ccctccctca ggcgtgatg ggctcttcat acggattcca atactctcct 9660  
 ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgccc tatgggcttc 9720  
 tcaatgaca cccgctgttt tgactcaacg gtactgaga atgacatccg tgttgaggag 9780  
 gcatcttacc aatgttgta cttggcccc gaagccagac aggccataag gtcgctcaca 9840  
 35 gagcggcttt acatcgggg cccctgact aattctaaag ggcagaactg cggctatcgc 9900  
 cggtgccggc cgagcgggtg actgacgacc agctgcggta ataccctcac atgttatttg 9960  
 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacgatgct cgtatgcgga 10020  
 gacgaccttg tcgttatctg tgaagcggc gggacccaa aggacgaggc gagcctacgg 10080  
 gccttcacgg aggttatgac tagatactct gcccccttg gggacccgccc caaaccagaa 10140  
 40 tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcatct 10200  
 ggcaaaaggg tgtactatct caccogtgac cccaccaccc cccttgccg ggctgcgtgg 10260  
 gagacagcta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgccc 10320  
 acctgtggg caaggatgat cctgatgact catttcttct ccaccttct agctcaggaa 10380  
 caacttga aaagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440  
 gacctacctc agatcattca acgactccat ggccttagcg cattttctact ccatagttac 10500  
 45 tctccaggtg agatcaatag ggtggcttca tgccctcagga aacttgggg accgcccctg 10560  
 cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620  
 ctgtccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680  
 ccaatcccg ctgcgtccca gttggattta tccagctggg tcgttgctgg ttacagcggg 10740  
 ggagacatat atcacagcct gtctcgtgcc cgaccccgt ggttcatgtg gtgcctactc 10800  
 ctactttctg taggggtagg catctatcta ctccccaa accgatgacgg gagctaaaca 10860  
 50 ctccaggcca ataggccatc ctgttttttt cctctttttt tttctttttt tttctttttt 10920  
 tttttttttt tttttttttt tctttttttt tctctttttt tttctttttt tttctttttt 10980  
 tggctccatc ttagecctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040  
 agagagtgtc gatactggcc tctctgcaga tcaagt 11076

<210> 8  
 <211> 8001  
 <212> DNA  
 <213> Hepatitis C Virus

<400> 8  
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60  
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcctgtag cctccaggac 120  
 ccccccctccc gggagagcca tagtggtctg cggaaaccgt gagtacaccg gaattgccag 180  
 gacgaccggg tcctttcttg gatcaaccgg ctcaatgect ggagatttgg gcgtgcccc 240  
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtaact gcctgatagg 300  
 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360  
 ctcaaaagaaa aaccaaacgt aacaccaacg ggcgcgcat gattgaacaa gatggattgc 420  
 acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480  
 caatcggtct ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540  
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgagga gcgcggctat 600  
 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttctc actgaagcgg 660  
 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccctg 720  
 ctctgcgcca gaaagtatcc atcatggctg atgcaatgcg gcggtgcat acgcttgatc 780  
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840  
 tggaaagccg tcttctcgat caggatgac tggacgaaga gcacgaggg ctcgcgccag 900  
 ccgaactggt cgccaggctc aaggcgcgca tgcccgcagg cgaggatctc gtcgtgacct 960  
 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020  
 actgtggccg gctgggtgtg gcggagccgt atcaggacat agcgttggct acccgtgata 1080  
 ttgtgaaga gcttggcgcc gaattgggtg accgcttctc cgtgctttac ggtatcgccg 1140  
 ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagttttaa 1200  
 cagaccacaa cgggttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260  
 aacgttactg ccgaagccg cttggaataa ggccggtgtg cgtttgteta tatgttattt 1320  
 tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtctctctg 1380  
 acgagcattc ctaggggctc tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440  
 gtgaaggtag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500  
 tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560  
 taagatacac ctgcaaaggc ggcacaaacc cagtgccacg ttgtgagttg gatagttgtg 1620  
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680  
 gtacccattg gtatgggata tgatctgggg cctcgggtga catgctttac atgtgtttag 1740  
 tcgagggttaa aaaacgtcta ggcctcccca accacgggga cgtgggtttc ctttgaaaaa 1800  
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860  
 tgcactcatc ctacgctcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg 1920  
 gtctccaccg caacacaatc ttctctggcg acctgcgtca atggcggtg ttggactgtc 1980  
 tatcatggtg ccggtctaaa gacccttgcc ggcccaaaag gcccaatcac ccaaatgtac 2040  
 accaatgtgg accagacact cgtcggctgg caagcgcccc ccggggcgcg ttccttgaca 2100  
 ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160  
 gtgcgcggcg ggggcgacag cagggggagc ctactctccc ccaggccccg ctcctacttg 2220  
 aagggtctct cgggcgggtc actgctctgc ccctcggggc acgctgtggg catcttctcg 2280  
 gctgccgtgt gcaccgaggg ggttgcaag gcgggtgact ttgtaccctg cgagtctatg 2340  
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400  
 acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460  
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgctcgt cgccgccacc 2520  
 ctagggttctg gggcgatatg gtctaaggca catggtatcg accctaacat cagaatcggg 2580  
 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640  
 gacgggtggt gctctggggg gcctatgac atcataatat gtgatgagt ccaactcaact 2700  
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760  
 cgactcgtcg tgctcgccac cgctacgctc ccgggatcgg tcaccggtcc acatccaaac 2820  
 atcgaggagg tggctctgtc cagcactgga gaaatcccc tttatggcaa agccatcccc 2880  
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940  
 gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000  
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaata 3060  
 acgggcttta ccggtgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120  
 gtcgacttca gcctggacct gaccttcacc attgagacga cgaccgtgcc acaagacggc 3180  
 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggtatgggcat ttacaggttt 3240  
 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtctat 3300  
 gacgcgggct gtgcttggtg cgaagctcac ccgcgcgaga cctcagttag gttgcgggct 3360  
 tacctaaaca caccaggggt gcccgctctg caggaccatc tggagttctg ggagagcgct 3420  
 tttaacaggg tcaccacat agacgcccct ttctgttccc agactaagca ggcaggagac 3480  
 aacttcccc acctggtagc ataccaggct acgggtgtgc ccagggtca ggctccacct 3540

	ccatcgtggg	accaaagtgt	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgcccctgc	tgtataggtc	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggttagcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
	attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccagacg	ggaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcccttacct	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgctcc	cgtggtggaa	tccaagtggc	ggaccatcga	agccttctgg	4020
	gcgaagcata	tgtggaatct	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
10	cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgctc	4140
	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttget	4200
	cctcccagcg	ctgcttctgc	tttcgttagg	gccggcatcg	ctggagcggc	tgttggcagc	4260
	atagccttg	ggaaggtgct	tgtggatatt	ttggcaggtt	atggagcagg	ggtggcaggc	4320
	gcgctcgtgg	cctttaaggt	catgagcggc	gagatgcctt	ccaccgagga	cctgggttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgacgcgata	4440
	ctgcgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggaatgaacc	gctgatagcg	4500
15	ttcgcttcgc	ggggttaacca	cgctcccccc	acgcactatg	tgccctgagag	cgacgctgca	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggtctcac	4620
	cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
	gattggatat	gcacggtgtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccc	4740
	cgattgcccg	gagtcctcct	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaacacc	ctgcccattg	ggggcacaga	tcaccggaca	tgtgaaaaac	4860
20	gggtcccatg	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
	attaacgcgt	acaccacggg	ccccctgcag	ccctccccgg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	cgacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tggttggggg	gcggttgca	aggtacgctc	cagcgtgcaa	acccctccta	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
	tgcgagcccg	aaccggatgt	agcagtgtct	acttccatgc	tcaccgaccc	ctccacatct	5280
25	acggcggaga	cggttaacgc	taggtcgccc	aggggatctc	ctccccctct	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactaccgc	tcagtactcc	5400
	ccggacgctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaaacatc	5460
	acccgcgttg	agtcagaaaa	taaggtagta	attttgact	ctttcgagcc	gctccaagcg	5520
	gaggaggtatg	agagggaagt	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
	cctcagcgca	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtctctg	5640
30	aaggaccccg	actacgtccc	tccagtggta	cacgggtgtc	cattgccggc	tgccaaggcc	5700
	cctccgatac	caacctcacg	gaggaaagag	acggttctcc	tgctcagaatc	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgtc	ggcgtctgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtcgt	actcctccat	gccccctctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gctagtggag	acgtcgtctg	ctgctcgatg	6000
	tcctcacat	ggacaggcgc	cctgatcagc	ccatgcgtcg	cggaggaaac	caagctgccc	6060
35	atcaatgcac	tgagcaactc	ttgtctcgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgacgcgcaa	gcctgcggca	gaagaaagtc	acctttgaca	gactgcaggt	cctggacgac	6180
	cactaccggg	acgtgtctca	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatattggc	6300
	tatgggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaaccacat	ccgctccgtg	6360
	tggaaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
40	gaggttttct	gcgtccaacc	agagaagggg	ggccgcgaagc	cagctcgctt	tatcgatttc	6480
	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
	cctcagggcg	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctgggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tgttttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtagcttgg	ccccgaagc	cagacaggcc	ataaggtcgc	tcacagagcg	gctttacatc	6780
45	ggggggcccc	tgactaatte	taaaggggcag	aactgcggct	atcgcgggtg	ccgcgcgagc	6840
	ggtgtactga	cgaccagctg	cggttaatacc	ctcacatggt	acttgaaggc	cgctgcggcc	6900
	tgctcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgoggggac	ccaagaggac	gaggcgagcc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccaaac	cagaatacga	cttgaggttg	7080
50	ataacatcat	gctcctccaa	tgtnctcagtc	gcgcacgatg	catctggcaa	aagggtgtac	7140
	tatctcacc	gtgacccca	caccccctt	gcgcgggtcg	cgtgggagac	agctagacac	7200
	actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgccacactt	gtgggcaagg	7260
	atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560  
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620  
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680  
agcctgtctc gtgcccgacc ccgctgggtc atgtgggtgcc tactcctact ttctgtaggg 7740  
gtaggcattct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctggt tttttccctt tttttttttc tttttttttt tttttttttt tttttttttt 7860  
ttttctcctt tttttttcct cttttttttcc ttttctttcc tttggtggct ccatcttagc 7920  
cctagtcacg gctagctgtg aaagggtccgt gagccgcttg actgcagaga gtgctgatac 7980  
10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 9  
 <211> 11076  
 <212> DNA  
 <213> Hepatitis C Virus

5

&lt;400&gt; 9

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccc gaattgccag 180
gacgacgggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaagaaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcagggtc tcgcggcgcct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaaggggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatcgg gcggtgcat acgcttgatc 780
cggctacctg cccattcgac caccagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcacag 900
ccgaactgtt cgcacggctc aaggcgcgca tggccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtcttcc tgagttaaaa 1200
cagaccacaa cgttttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tggcgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctc ttccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
gtgaaggaa gcagttctct ggaagcttct tgaagacaaa caacgtctgt agcgacctt 1500
tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtggcacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccactt gtatgggata tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tctcgtgtgg agtttacctg 1920
ttgcgcgcga ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccagagg tagggcctgg 2040
gctcagcccg ggtaccctcg gccctctat ggcaatgagg gcttgggttg ggcaggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggaaccccg gcgtaggttg 2160
cgcaatttgg gtaaggctcat cgataccctc acgtgcggct tcgccgatct catggggtac 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgctcgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc ctttctatc 2340
ttccttttgg ctttgtgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccgggtgc gtgcectgcg ttcgggagaa caactctcc 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtgcgccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggctac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagg tgtcgtggat atggtggcgg gggccattg gggagtccta 2880
gcgggccttg cctactatc catggtgggg aactgggcta aggttctgat tgtgatgta 2940
ctccttgccg gcgttgacgg gggaaacctat gtgacgggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttaccccggg tcacccaga aaatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
gccagctgca gcccacgca cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccgcggcc gtgcggtatc 3300
gtacccgcgg cgcaggtgtg tggccagtg tactgttca cccaagccc tgcgtggtg 3360
gggacgacgg accggttcgg cgtccctacg tacagtggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gcccgcgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc cccctgtgta acatcggggg gatcggcaat 3540

```

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55

```

aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
tgtggttcgg ggccttggtt gacaccacaga tgcttggtcc actaccacata caggcctttgg 3660
cactaccctt gcaactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcggttgtaa cctggaggac 3780
agggacagat cagagcttag ccgctgtctg ctgtctacaa cggagtggca ggtattgccc 3840
tgttccttca ccaccctacc ggcctgtctc actggtttga tccatctcca tcagaacgtc 3900
gtggacgtac aatacctgta cgttataggg tcggcgggtg tctcctttgc aatcaaatgg 3960
gagtatgtcc tgttgctctt ccttctctct gcgagcgcgc gcgtctgtgc ctgcttgtgg 4020
atgatgtgct tgatagctca agctgaggcc gccctagaga acctggtggt cctcaacgcg 4080
gcatccgtgg ccggggcgca tggcattctc tcttctctcg tgttctctct tgctgcctgg 4140
tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200
ctcctgtccc tgctggcgtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
tcgtgaggag gcgcggtttt cgtaggctctg atactcttga ccttgtcacc gactataag 4320
ctgttctctcg ctaggctcat atggtggtta caatatTTTA tcaccagggc cgaggcacac 4380
ttgcaagtgt ggatcccccc cctcaacgtt cggggggggc gcgatgcgtt catcctctct 4440
acgtgcgcga tccaccacga gctaatcttt accatcacca aaatcttgct cgccatactc 4500
gggtccactca tgggtgctcca gctggtgctg accaaagtgc cgtactctgt gcgcgcacac 4560
gggtctattc gtgcatgcat gctggtgctg aaggttgctg ggggtcatta tgtccaaatg 4620
gctctcatga agttggcgcg actgacaggt acgtacgttt atgacctctt caccctactg 4680
cgggactggg cccacgcggg cctacgagac cttgcgggtg cagttgagcc cgtcgtcttc 4740
tctgatatgg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atcttgggcc tgcccgtctc cgcccgcagg gggagggaga tacatctggg accggcagac 4860
agccttgaag ggcagggggt gcgactcttc gcgcctatta cggcctactc ccaacagacg 4920
cgaggcctac ttggctgcat cctacatagc ctacacaggcc gggacaggaa ccaggctcag 4980
ggggaggctc aagtggctct caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtggttga ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggecca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gcccccggg 5160
gcgcgttcct tgacaccatg cacctgcggc agctcggacc tttacttggt cagcaggcat 5220
gcgatgtca ttccggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5280
cccgctctct acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcacgct 5340
gtgggcatct ttccggctgc cgtgtgcacc cgagggggtg cgaaggcggt ggactttgta 5400
cccgctcagt ctatggaaac cactatgcgg tccccggtct tcacggacaa ctctgcccc 5460
ccggccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gcatgcagcc caagggtata aggtgcttgt cctgaacccg 5580
tccgtcgcgg ccaccctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640
aacatcagaa tcggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacct 5700
ggcaagtctc ttgccgacgg tgggtgtctc gggggcgccct atgacatcat aatatgtgat 5760
gagtgcact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
gagacggctg gagcgcgact cgtcgtgtct gccaccgcta cgcctccggg atcggtcacc 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctgggaaat ccccttttat 5940
ggcaaagcca tccccatcga gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
aagaagaaat gtgatgagct cgccgcgaag ctgtccggcc tcggactcaa tgctgtagca 6060
tattaccggg gcttgatgtt cttaccgggt gacttcgact cagtgatcga ctgcaatata 6180
acggacgtct taatgacggg cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
tgtgtcaccc agcgggtgtc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
gtgccacaag ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttctcgggt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca ggggtgccc tctgccagga ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggtacagg gtctcatag 6600
gctcaggctc caccctcatc gtgggaccaa atgtggaagt gtctcatag cgtaaaagcct 6660
acgctgcagc ggccaacgcc cctgctgtat aggtcgggag ccgttcaaaa cgaggttact 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
acgagcacct ggtgtgctgt aggcggagtc ctgacagctc tggccgcgta ttgcctgaca 6840
acaggcagcg tggctattgt gggcaggatc atcttgcctg gaaagccggc acacctccct 6900
gacagggaaag tctttaccg ggagtctgat aggtcgggag agtgcgccct acacctccct 6960
tacatcgaaac agggaaatgc gctcgccgaa caattcaaac agaaggcaat cgggttgctg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaaatcaa gtggcgagc 7080
atcgaaacct tctgggcgaa gcatatgtgg aatttcatca gcgggatata atatttagca 7140
ggcttgtcca ctctgcctgg caaccocgcg atagcateac tgatggcatt cacagcctct 7200
atcaccagcc cgctcacacc ccaacatacc ctectgttta acatcctggg gggatgggtg 7260
gccgccaaac ttgctcctcc cagcgctgct tctgctttcg tagcgccggg catcgctgga 7320
gcggctgttg gcagcatagg ccttgggaaag gtgcttgtgg atattttggc aggttatgga 7380
gcagggggtg caggcgcgct cgtggccttt aaggctatga gcggcgagat gccctccacc 7440
gaggacctgg ttaacctact ccctgctatc ctctcccctg gcgccttagt cgtcggggtc 7500

```

5 gtgtgcgcag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560  
 aaccggctga tagcggttcgc ttcgggggtt aaccacgtct ccccccacgca ctatgtgcct 7620  
 gagagcgacg ctgcagcacg tgtcactcag atcctctcta gtcttaccat cactcagctg 7680  
 ctgaagaggc ttcaccagtg gatcaacgag gactgctcca cgccatgctc cggtcgtggtg 7740  
 ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccaag 7800  
 tccaagctcc tgccgcgatt gccgggagtc cccttcttct catgtcaacy tgggtacaag 7860  
 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggggc acagatcacc 7920  
 ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980  
 catggaacat tcccattaa cgcgtacacc acgggcccct gcacgcccct cccggcgcca 8040  
 10 aattattcta gggcgctgtg cgtgggtggt gctgaggagt acgtggaggt tacgctgggtg 8100  
 ggggatttcc actacgtgac gggcatgacc actgacgacg taaagtgcc gtgtcaggtt 8160  
 ccggcccccg aattcttcac agaagtggat ggggtgcggt tgcaacaggta cgctccagcg 8220  
 tgcaaaacccc tcttacggga ggaggtcaca ttcctggctg ggtcactca atacctggtt 8280  
 gggtcacagc tcccatgcga gccgaacccg gatgtagcag tgctcacttc catgctcacc 8340  
 gacccctccc acattacggc ggagacggct aagcgtaggc tggccagggg atctcctccc 8400  
 15 cccttgggcca gctcatcagc tagccagctg tctgcgcctt ccttgaaggc aacatgcact 8460  
 acccgctcatg actccccgga cgtgacctc atcgaggcca acctcctgtg gcggcaggag 8520  
 atggcgcgga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580  
 gagccgctcc aagcggagga ggatgagagg gaagtatccg ttcggcgcca gatcctgcgg 8640  
 aggtccagga aattccctcg agcgtatgcc atatgggcac gcccgatta caaccctcca 8700  
 ctgttagagt cctggaagga cccggactac gtccctccag tggtaacagg gtgtccattg 8760  
 20 ccgcctgcca agggccctcc gataccactc tcacggagga agaggacggt tgcctgtca 8820  
 cgaatctacc tggtcttctg cttggcgagg ctcgccacag agaccttcgg cagctccgaa 8880  
 tcgtcgcccg tcgacagcgg cacggcaacg gcctctcctg accagccctc cgacgacggc 8940  
 gacgcgggat ccgacgttga gtcgtactcc tccatgcccc cccttgaggg ggagccgggg 9000  
 gatcccgatc tcagcgacgg gtcttggctc accgtaagcg agggagctag tgaggacgtc 9060  
 gtctgctgct ccatgtccta cacatggaca ggcgcctga tcacgcatg cgtgcggaag 9120  
 25 gaaaccaagc tgcccatcaa tgactgagc aactctttgc tccgtacca caacttggtc 9180  
 tatgtacaaa catctcgag cgcaagcctg cggcagaaga aggtcacctt tgacagactg 9240  
 caggtcctgg acgaccacta cccggagctg ctcaaggaga tgaaggcgaa ggcgtccaca 9300  
 gttaaggcta aacttctatc cgtggaggaa gcctgtaagc tgacgcccc acattcggcc 9360  
 agatctaaat ttggctatgg ggcaaaggac gtccggaacc tatccagcaa ggcctgtaac 9420  
 30 cacatccgct ccgtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480  
 atcatggcaa aaaatgaggt tttctgcgtc caaccagaga aggggggccc caagccagct 9540  
 cgccttatcg tattccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600  
 gtggcttcca ccctccctca ggccgtgatg ggctcttcat acggattcca atactctcct 9660  
 ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgccc tatgggcttc 9720  
 gcatatgaca cccgctgttt tgactcaacg gtcactgaga atgacatccg tgttgaggag 9780  
 35 tcaatctacc aatgttga cttggcccc gaagccagac aggcataag gtcgctcaca 9840  
 gacggtcttt acatcggggg cccctgact aattctaaag ggcagaactg cggctatcgc 9900  
 cgggtgccgg cgagcgtgt actgacgacc agctgcggta ataccctcac atgttacttg 9960  
 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacgatgct cgtatgcgga 10020  
 gacgaccttg tcgttatctg tgaagcgcg gggacccaag aggcagaggc gagcctacgg 10080  
 gccttcacgg aggtatgac tagatactct gccccccctg gggacccgccc caaaccagaa 10140  
 40 tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcatct 10200  
 ggcaaaaggg tgtactatct caccctgac cccaccaccc cccttgccg ggctgcgtgg 10260  
 gagacagcta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgccc 10320  
 acctgtggtg caaggatgat cctgatgact catttcttct ccatccttct agctcaggaa 10380  
 caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440  
 gacctacctc agatcattca acgactccat ggccttagcg cattttctact ccatagttac 10500  
 45 tctccagggt agatcaatag ggtggcttca tgctcaggga aacttggggg accgccttg 10560  
 cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620  
 gctgcccatt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680  
 ccaatccccg ctgctgcccc gttggattta tccagctggt tcgttgctgg ttacagcggg 10740  
 ggagacatat atcacagcct gtctcgtgcc cgaccccgct ggttcatgtg gtgcctactc 10800  
 ctactttctg taggggtagg catctatcta ctccccaccc gatgaacggg gagctaaaca 10860  
 50 ctccaggcca ataggccatc ctgttttttt cctttttttt tttctttttt tttctttttt 10920  
 tttttttttt tttttttttt tccttttttt ttctcttttt tttctttttt tttctttttt 10980  
 tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040  
 agagagtgtc gatactggcc tctctgcaga tcaagt 11076

<210> 10  
 <211> 8001  
 <212> DNA  
 <213> Hepatitis C Virus

<400> 10  
 gccagccccc gattgggggc gacactccac catagatcac tccccgtga ggaactactg 60  
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120  
 cccccctccc gggagagcca tagtggtctg cggaaacggg tagtacaccg gaattgccag 180  
 gacgacccgg tcccttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240  
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300  
 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatccctaac 360  
 ctcaaaagaaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420  
 aacgaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480  
 caatcggtctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggcgcc ccggttcttt 540  
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggtat 600  
 cgtggctggc cagcaggggc gttccttgcc cagctgtgct cgacgttgc actgaagcgg 660  
 gaagggactg gctgctattg ggcaagtgc cggggcaggga tctcctgtca tctcaccttg 720  
 ctctgcccga gaaagtatcc atcatggctg atgcaatgcg gcggtgcac acgcttgatc 780  
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840  
 tggaaagccg tcttgcgat caggatgac tggacgaaga gcatcagggg ctgcgcagc 900  
 ccgaactggt cgccaggctc aaggcgccga tgcgcgacgg cgaggatctc gtcgtgaccc 960  
 atggcgatgc ctgcttgccc aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020  
 actgtggccc gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080  
 ttgctgaaga gcttggcgcc gaattgggtg accgcttctc cgtgctttac ggtatcgccg 1140  
 ctccccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200  
 cagaccacaa cggtttccct ctgacgggat caattccgcc cctctccctc cccccccct 1260  
 aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320  
 tccaccatat tgccgtcttt tggcaatgtg agggcccggg aaactggccc tgtcttcttg 1380  
 acgagcattc ctgagggtct tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440  
 gtgaaggagc cagttcctct tgaagacaaa caacgtctgt agcgaccctt 1500  
 tgcaggcagc ggaaccccc accctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560  
 taagatacac ctgcaaaagg gcacaaaccc cagtgcacgg ttgtgagttg gatagttgtg 1620  
 gaaagagtca aatggctctc ctcaagcgtg ttcaacaagg ggctgaaggga tgcccagaag 1680  
 gtaccctatt gtatgggata tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740  
 tcgaggttaa aaaacgtcta gggcccccca accacgggga cgtgggtttc ctttgaataa 1800  
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860  
 tgcacatcac ctgacctcac aggcgggac aggaaccagg tcgaggggga ggtccaagtg 1920  
 gtctccaccc caacacaatc tttctggcg accctgcgtc atggcgtgtg ttggactgtc 1980  
 tatcatggtg ccgggtcaaa gacccttgcc ggcccaagg gcccaatcac ccaaatgtac 2040  
 accaatgtg accagacct cgtcggtgg caagcgcccc ccggggcgcg ttccttgaca 2100  
 ccatgcacct gcggcagctc ggacctttac ttggtcagca ggcatgccga tgtcattccg 2160  
 gtgcgcggcg gggggcagac cagggggagc ctactctccc ccaggcccgt ctccctactt 2220  
 aagggtctct cgggcggtcc actgctctgc ccctcggggc acgctgtggg catcttctcg 2280  
 gctgcccgtg gcacccgagg ggttgcaag gcggtggact ttgtaccgt cgagtctatg 2340  
 ggaaccacta tgcgggtccc ggtcttcac gacaactcgt cccctccggc cgtaccgcag 2400  
 acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460  
 gctgcgatg caggccaagg gtataagggt ctgttccctg acccgctcgt cgccgccacc 2520  
 ctagggttctg gggcgatat gtctaaggca catggtatcg accctaacat cagaatcggg 2580  
 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640  
 gacgggtggt gctctggggg cgcttatgac atcataatat gtgatgagtg ccaactcaact 2700  
 gactcgacca ctatcctggg catcggcaca gtccctggacc aagcggagac ggctggagcg 2760  
 cgactcgtcg tgtctgccac cgctacgcct ccgggatcgg tcaccggtcc acatccaaac 2820  
 atcgaggagg ttggtctgtc cagcactgga gaaatcccc tttatggcaa agccatcccc 2880  
 atcgagacca tcaagggggg gaggcaacct attttctgcc attccaagaa gaaatgtgat 2940  
 gagctcgccg cgaagctgtc cgccctcgga ctcaatgctg tagcatatta ccggggcctt 3000  
 gatgtatccg tcataccac tagcggagac gtcattgtcg tagcaacgga cgtctaatg 3060  
 acgggcttta ccggcgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120  
 gtcgacttca gcctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacggc 3180  
 gtgtcacgct cgcagcgccg aggcaggact ggtaggggca ggaatgggcat ttacaggttt 3240  
 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300  
 gacgcggtct gtgcttggta ctgactcacg cccgcccaga cctcagttag gttgcgggct 3360  
 tacctaaaca caccagggtt gccggtctgc caggaccatc tggagttctg ggagagcgtc 3420  
 tttacaggcc tcaccacat agacgcccac ttcttgtccc agactaagca ggcaggagac 3480  
 aacttcccc acctggtagc ataccaggct acggtgtgcg ccagggtcca ggctccacct 3540

5

10

15

20

25

30

35

40

45

50

```

ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacggggcca 3600
acgcccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acaccccata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtagggc gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
attgtgggca ggatcatctt gtccggaaag ccggccatca tccccgacag ggaagtccct 3840
taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caagcggagg ctgctgtccc cgtggtggaa tccaagtggc ggaccatcga agccttctgg 4020
gcgaagcata tgtggaattt catcagcggg atacaatatt tagcaggctt gtccactctg 4080
cctggcaacc ccgcgatagc atcactgatg gcattcacag cctctatcac cagcccgctc 4140
accaaccaac ataccctcct gtttaacatc ctggggggat ggggtggccg ccaacttgct 4200
cctcccagcg ctgcttctgc tttcgtagc gccggcatcg ctggagcggc tgttggcagc 4260
ataggccttg ggaaggtgct tttcgtagc gccggcatcg ctggagcggc tgttggcagc 4320
gcgctcgtgg cctttaaggt catgagcggc gaggttgccc ccaccgagga cctgggttaac 4380
ctactccctg ctatcctctc ccctggcgcc ctagtctgct gggctgctgt cgacgcgata 4440
ctgcgtcggc acgtggggcc aggggagggg gctgtgcagt ggaatgaacc gctgatagcg 4500
ttcgcttcgc ggggtaacca cgtctcccc ccgcaactatg tgctgagag cgacgctgca 4560
gcacgagtca ctcatctct ctctagtctt accatcactc agctgctgaa gaggtctcac 4620
cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaa gctcctgccc 4680
gattggatga gcacgggtgtt gactgatttc aagacctggc tccagtccaa gctcctgccc 4740
cgattgcggg gagtccctct ctctcatgt caacgtgggt acaagggagt ctggcggggc 4800
gacggcatca tgcaaacacc ctgcccattg ggagcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggatcgtggg gcctaggacc ttagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccaagg cccttcgacg cctcccccg ccgcaaatla ttttagggcg 4980
ctgtggcggg taggtgctga ggagtacgtg gaggttacgc ggggtggggg tttccactac 5040
gtgacgggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccgaaattc 5100
ttcacagaag tggatggggt gcggttgca aggtacgctc cagcgtgcaa acccctcta 5160
cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctcccc 5220
tgcgagcctg aaccggatgt agcagtgtct acttccatgc tcaccgaccc ctcccacatt 5280
acggcggaga cggctaagcg taggtggccc aggggatctc cccccctt ggccagctca 5340
tcagctagcc agctgtctgc gccttctctg aaggcaacat gcactaccgg tcatgactcc 5400
ccggacgctg acctcatcga ggccaacctc ctgtggcggc agggatggg cgggaaacatc 5460
accgcgtgg agtcagaaaa taaggtagta attttgact ctttcgagcc gctccaagcg 5520
gaggaggatg agaggggagt atccgttccg gcggagatcc tgccggaggtc caggaaattc 5580
cctcgagcga tgcccatatg ggacgcccgg gattacaacc ctccactgtt agagtccctg 5640
aaggacccgg actacgtccc tccagtgtga caccgggtgtc cattgccgcc tgccaaggcc 5700
cctccgatac cacttccagc gaggaaaggg acggttgctc tgcagaatc taccgtgtct 5760
tctgccttgg cggagctcgc cagagagacc ttcggcagct ccgaatcgtc ggccgtcgac 5820
agcggcacgg caacggcctc tccagaccag ccctccgacg acggcgacgc gggatccgac 5880
gttgagtcgt actcctccat gccccctt gagggggagc cgggggatcc cgaatcagc 5940
gacgggtctt ggtctaccgt aagcgaggag gctagtggg acgtcgtctg ctgctcgatg 6000
tcctacacat ggacaggcgc cctgatcagc ccatgcgctg cggaggaaac caagctcccc 6060
atcaatgcac tgagcaactc tttgtccgt caccacaact tggcttatgc tacaacatct 6120
cgacagcgca acctgcggca gaagaaggtc accttgaca gactgcaggt cctggacgac 6180
cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
ctatccgtgg aggaagcctg taagctgacg cccccacatt cggccagatc taaatttggc 6300
tatggggcaa aggacgtccg gaacctatcc agcaaggccc ttaaccacat ccgctccgtg 6360
tggaaggact tgcgtggaag cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgcct tatcgtatc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
cctcaggccc tgatgggctc ttcatacgga ttccaatact ctctgggaca gcgggtcgag 6600
ttcctgggtg atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacaccgcg 6660
tgttttgact caacggtcac tgagaatgac atccgtgttg agggatcaat ctaccaatgt 6720
tgtagcttgg cccccgaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
gggggcccc tgactaatc taaagggcag aactgcggct atcgcgggtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cggttaatac ctacatgtt actgaaggc cgctgcggcc 6900
tgctgagctg cgaagctcca ggactgcagc atgctcgtat gcggagacga ccttgcgtt 6960
atctgtgaaa gcgcggggac ccaagaggac gaggcgagcc tacgggctt cacggaggct 7020
atgactagat actctgcccc ccctggggac ccgcccaaac cagaatacga cttggagttg 7080
ataacatcat gctcctccaa tgtgtcagtc gcgcacgatg catctggcaa aagggtgtac 7140
tatctcaccg gtgaccccc caccctctt gcgcgggctg cgtgggagac agctagacac 7200
actccagtc atctctggct aggcaacatc atcatgtatg cgccacett gtgggcaagg 7260
atgactctga tgactcatt ctctccatc ctctagctc aggaacaact tgaaaaagcc 7320
ctagattgtc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380
attcaacgac tccatggcct tagcgcatct tcaactcata gttactctcc aggtgagatc 7440
aatagggtgg ctctcatgct caggaaactt ggggtaccgc cctgcgagt ctggagacat 7500

```

55

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560  
 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620  
 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680  
 agcctgtctc gtgcccgcgc ccgctgggtc atgtggtgcc tactcctact ttctgtaggg 7740  
 gtaggcacat atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
 ccatcctggt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860  
 ttttctcctt tttttttcct ctttttttcc ttttctttcc tttggtggct ccatcttagc 7920  
 cctagtcacg gctagctgtg aaagggtccgt gagccgcttg actgcagaga gtgctgatac 7980  
 10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 11  
 <211> 11076  
 <212> DNA  
 <213> Hepatitis C Virus

5

<400> 11

10

15

20

25

30

35

40

45

50

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttggtgc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcaagtgcc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggtctacct cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgcgat caggatgac tggacgaaga gcatcagggg ctccgcccag 900
ccgaactgtt cgccaggtcc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgcg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg gctgggtgtg gcggaaccgt atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggtcg accgcttctt cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200
cagaccacaa cggtttccct ctacggggat caattccgcc cctctccctc cccccccctt 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccggg aacctggccc tgtcttcttg 1380
acgagcattc ctagggtgtt ttccctctc gccaaaggaa tgcaaggtct gttgaatgct 1440
gtgaaggaa gcaagcttct tgaagcaaaa caacgtctgt agcgaccctt 1500
tgaggcagc ggaaccccc accctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaagg gccacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccatt gtatgggata gctcgtggg cctcgttgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta gggccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cagcataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacy taacaccaac 1860
cgccgcccac aggcagtcac gttcccgccc ggtggtcaga tctcgtgtgg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctc 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccagagg tagggcctgg 2040
gctcagcccc ggtacccctg gccctctat ggcaatgagg gcttgggggt ggcaggatgg 2100
ctcctgtcac cccgtggctc tggggcccc cgacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggctcat cgataccctc acgtgcggct tcccgatct catggggtac 2220
attccgctcg tcggcggccc cctagggggc gctgccagg ccctggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgct ctttctctac 2340
ttccttttgg ctttctgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cagcaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccggtgct gtgccctgct ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga ttgtctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctct gtcgccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtaacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagg tgcgtgggat atggtggcgg gggcccatgt gggagtccta 2880
gcgggccttg cctactatcc catgggtggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgcgg gcgttgacgg gggaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcaaccggg tcatcccaga aatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
gccagctgca gcccacatga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
tcacacagct cgacacagag gccttattgt ttgcaactac caccocggcc gtgcggatc 3300
gtacccgcgg cgcaggtgtg tggctccagt tactgcttca cccaagccc tgcgtggtg 3360
gggacgaccg accggttcgg cgtccctacg tacagtggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgccgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc cccctgtgta acatcggggg gatcggaat 3540

```

55

5

10

15

20

25

30

35

40

45

50

55

```

aaaaccttga cctgccccac ggactgtctc cggaagcacc ccgaggccac ttacaccaag 3600
tgtgggttcgg ggccttgggtt gacaccacaga tgcttgggtcc actaccata caggcttttg 3660
cactaccctt gcactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tcgaagccgc atgcaatttg actcaggag agcgtgttaa cctggaggac 3780
agggacagat cagagcttag ccgctgtctg ctgtctacaa cggagtggca ggtattgccc 3840
tgttccctca ccaccctacc ggcctgtctc actggtttga tccatctcca tcagaaacgtc 3900
gtggacgtac aatacctgta cggtataggg tcggcgggtg tctcctttgc aatcaaatgg 3960
gagtatgtcc tgttgctctt ccttctctcg gcggacgcgc gcgtctgtgc ctgcttgtgg 4020
atgatgtctg tgaatagctc agctgaggcc gccctagaga acctggtggt cctcaacgcg 4080
gcatccgtgg ccggggcgca ccttctctcg tgcttctctg tgctgctctg 4140
tacatcaagg gcaggctggg ccttggggcg gcataatgccc tctacggcgt atggccgcta 4200
ctcctgtctc tgctggcggtt accaccacga gcatacgcga tggaccggga gatggcagca 4260
tcgtgaggag gcgcggtttt cgtaggctcg atactcttga cctgtcacc gcactataag 4320
ctgttctctc ctaggctcat atgggtggtt caatatttta tcaccagggc cgaggcacac 4380
ttgcaagtgt ggatccccc cctcaacggt cgggggggcc gcgatgcgt catcctctc 4440
acgtgcgaga tccaccacga gctaattctt accatcacca aaatcttgc cgccatactc 4500
ggtccactca tgggtgtcca ggctgtgata accaaagtgc cgtacttctg gcgcgcacac 4560
gggctcatcc gtgcatgcat gctggtgcgg aaggttgcgt ggggtcatta tgtccaaatg 4620
gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccctactg 4680
cgggactggg ccacacgcgg cctacgagac ctgtcggttg cagttgagcc cgtcgtcttc 4740
tctgatatgg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atctctggcc tgcccgtctc cgcgcgcagg gggagggaga tacatctggg accggcagac 4860
agccttgaag ggcagggtg catcactagc cgcctctatta cggcctactc ccaacagacg 4920
cgaggcctac ttggctgcat caccgcaaca caatctttcc tggcgacctg cgtcaatggc 4980
ggggaggtcc aagtggtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtgttggg ctgtctatac tgggtccggc tcaaagacc ttgccggccc aaaggggcca 5100
atcacccaaa tgtgaccag gacctcgtcg gctggcaagc gcccggggg 5160
gcgcttctc tgacaccatg cacctgcggc agctcggacc ttacttggg cagcaggcat 5220
cccgatgta ttcgggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5280
cccgtctctc acttgaagg cctctcgggc ggtccactgc tctgcccctc ggggcacgct 5340
gtgggcatct ttcgggctgc cgtgtgcacc cgagggttg cgaaggcggt ggactttgta 5400
ccgtcgagt ctatgggaac cactatgcgg tccccggtct tcacggacaa ctctgtcccct 5460
ccgcccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gtatgcaggc caagggtata aggtgcttgt cctgaaccgg 5580
tccgtcgcgc ccaccctagg tttcggggcg tataatgtcta aggcacatgg tatcgaccct 5640
aacatcagaa tcggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
ggcaagtttc ttgcccagcg tgggtgtctc gggggcgccct atgacatcat aatatgtgat 5760
gagtgcacct caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
gagacggctg gagcgcgact cgtcgtgtcc gccaccgcta cgcctccggg atcggtcacc 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaaaagcca tccccatcga gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
aagaagaat gtgatgagct cgcgcggaag ctgtccggcc tcggactcaa tgctgtagca 6060
tattaccggg gcttcatgtt atccgtcata ccaactagcg gagacgtcat tctcgtagca 6120
acggacgctc taatgacggg ctttaccggc gacttcgact cagtcatcga ctgcaatata 6180
tgtgtcaccg agacagtcca cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
gtgccacaag acgctggtgc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttctcgggt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcgcgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca ggggtgccg tctgccagga ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggtacggt gtgcgccagg 6600
gctcaggctc caccctccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660
acgctgcacg ggccaacgcc cctgtgttat aggtcgggag ccgttcaaaa cgaggttact 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
acgagcacct ggtgtctggt aggcggagtc ctacgagctc tggccgcgta ttgcttgaca 6840
acaggcagcg tggctattgt gggcgagatc atctgttccg gaaagccggc catcattccc 6900
gacagggaag tcttttaccg ggagttcgat gagatggaag agtgcgcctc acacctccct 6960
tacatcgaac aggggaatga gctcgccgaa caattcaaac agaaggcaat cgggttgcgt 7020
caaacagcca ccaagcaagc ggaggtgctc gctcccgtgg tggaaatcaa gtggcggacc 7080
atcgaagcct tctgggcgaa gcataatgtg aatttcata gcgggataca atatttagca 7140
ggcttgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
atcaccagcc cgtcaccac ccaacatacc ctctgttcta acatcctggg gggatgggtg 7260
gcgcccacac tctctctctc cagcgctgtc tctgcttttc taggcgcgg catcgctgga 7320
gcggctgttg gcagcatagg ccttgggaag gtgcttgttg atattttggc aggttatgga 7380
gcaggggttg caggcgcgct cgtggccttt aaggtcatga gcggcgagtt gccctccacc 7440
gaggacctgg ttaacctact cctgtctatc ctctcccctg gcgccctagt cgtcggggtc 7500

```

gtgtgacgag cgatactgag tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560  
 aaccggctga tagcgttcgc ttcgcggggt aaccacgtct ccccccacgca ctatgtgcct 7620  
 gagagcgacg ctgcagcagc agtcactcag atcctctcta gtcttaccat cactcagctg 7680  
 5 ctgaagaggc ttcaccagtg gatcaacgag gactgtctca cgccatgctc cggctcgtgg 7740  
 ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800  
 tccaagctcc tgccgagatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860  
 ggagtctggc gggcgacgag catcatgcaa accacctgcc catgtggagc acagatcacc 7920  
 ggacatgtga aaaaacgggtc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980  
 catggaacat tccccattaa cgcgtacacc acgggcccct gcacgccctc cccggcgcca 8040  
 10 aattattcta gggcgctgtg gcggttaggt gctgaggagt acgtggaggt tacgcggttg 8100  
 ggggatttcc actacgtgac gggcatgacc actgacaacg taaagtgccg gtgtcagggt 8160  
 ccggcccccg aattcttcac agaagtggat ggggtgcggt tgcacaggtg cgctccagcg 8220  
 tgcaaacccc tccatcgga ggaggtcaca ttcctggctg ggctcaatca atacctggtt 8280  
 gggtcacgac gacctgacg gcctgaaccg gatgtagcag tgctcacttc catgctcacc 8340  
 gacccctccc acattacggc ggagacggct aagcgtaggc tggccagggg atctcccccc 8400  
 15 cccttgccca gctcatcagc tagccagctg tctgcgcctt ccttgaaggc aacatgcact 8460  
 acccgctcat actccccgga cgtgacctc atcgaggcca acctcctgtg gcggcaggag 8520  
 atggcgcgga acatcacccg cgtggagtc gaaaataagg tagtaatttt ggactcttcc 8580  
 gagccgctcc aagcggagga ggatgagagg ggagtatccg ttcggcgga gatcctgcgg 8640  
 aggtccagga aattccctcg agcgtgccc atatgggac gcccgatta caacctcca 8700  
 ctgttagagt cctggaagga cccggactac gtccctccag tggtagacgg gtgtccattg 8760  
 20 ccggctgcca agggccctcc gatccacct tcacggagga agaggacggt tgtcctgtca 8820  
 gaatctaccg tgtcttctgc cttggcgagg ctgcacacag agaccttcgg cagctccgaa 8880  
 tcgtcgcccg tcgacagcgg cacggcaacg gcctctcctg accagccctc cgacgacggc 8940  
 gacgcgggat ccgacgttga gtcgtactcc tccatgcccc cccttgaggg ggagccgggg 9000  
 gatcccgat tcagcgacgg gtcttggctc accgtaagcg agggaggctag tgaggacgtc 9060  
 25 gctctgctgt cgatgtccta cacatggaca ggcgcctga tcacgccatg cgtcgggag 9120  
 gaaaccaagc tgcccatcaa tgcactgagc aactctttgc tccgtcacca caacttggtc 9180  
 tatgtacaaa catctcgag cgcaaacctg cggcagaaga aggtcacctt tgacagactg 9240  
 caggctcctg acgaccacta cgggacgtg ctcaaggaga tgaaggcga ggcgtccaca 9300  
 gttaaaggta aacttctatc cgtggaggaa gcctgttaagc tgacgcccc acattcggcc 9360  
 agatctaaat ttggctatgg ggcaaggac gtccggaacc tatccagcaa ggccgttaac 9420  
 30 cacatccgct ccgtgtgga ggacttgctg gaagacactg agacaccaat tgacaccacc 9480  
 atcatggcaa aaaatgaggt tttctgcgtc caaccagaga agggggggccg caagccagct 9540  
 cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600  
 gtggtctcca ccctccctca ggcggtgatg ggctcttcat acggattcca atactctcct 9660  
 ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgcc tatgggcttc 9720  
 gcatatgaca cccgctgttt tgactcaacg gtcactgaga atgacatccg tgttgaggag 9780  
 35 tcaatctacc aatgttgta cttggcccc cttgcccccc aggccataag gtcgctcaca 9840  
 gagcggttt acatcgggg cccctgact aattctaaag ggcagaactg cggctatcgc 9900  
 cgggtccgag cgagcggtgt actgacgacc agctgcggta ataccctcac atgttacttg 9960  
 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacgatgct cgtatgcgga 10020  
 gacgaccttg tcgttatctg tgaagcgcg gggacccaag aggcagagg gagcctacgg 10080  
 gccttcacgg aggttatgac tagatactct gccccccctg gggacccgcc caaaccagaa 10140  
 40 tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcactc 10200  
 ggcaaaaggg tgtactatct caccctgac cccaccaccc cccttgccg ggctgcgtgg 10260  
 gagacagcta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgcc 10320  
 acctgtggg caaggatgat cctgatgact catttcttct ccatccttct agctcaggaa 10380  
 caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440  
 gacctacctc agatcattca acgactccat ggccttagcg cattttcact ccatagttac 10500  
 45 tctccagggt agatcaatag ggtggcttca tgctcagga aacttgggt accgcccttg 10560  
 cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620  
 gctgccactt ttggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680  
 ccaatcccgg ctgctgcca gttggattta tccagctggg tcggtgctgg ttacagcggg 10740  
 ggagacatat atcacagcct gtctcgtgcc cgaccccgt ggttcatgtg gtgcctactc 10800  
 ctactttctg taggggtagg catctatcta ctcccaacc gatgaacggg gagctaaaca 10860  
 50 ctccaggcca ataggccatc ctgttttttt cctttttttt ttttcttttt ttttcttttt 10920  
 tttttttttt tttttttttc tctttttttt ttctcttttt ttctcttttt ttctcttttt 10980  
 tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040  
 agagagtgtc gatactggcc tctctgcaga tcaagt 11076

# **Patentansprüche**

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, dadurch gekennzeichnet,

daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen von einer handelsüblichen Hepatomazelllinie abstammen.

3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen aus einer Hepatomaprimärzellkultur gewonnen sind.

4. HCV-RNA-Konstrukt, dadurch gekennzeichnet,

daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt.

6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß die 3' NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT  
TTAGCTTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTG  
TTGCCCAGACTGGAGT

5 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GGAGC

10 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GCAGC

15 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
AGTCTCACTC TGTTGCCAG ACTGGAGT

20 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
AGTCTCACTC TGTTGCCAG ACTGGAGT

25 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
GAGTCTCACT CTGTTGCCA GACTGGAGT

30 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
GTCTCACTCT GTTGCCAG ACTGGAGT

35 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTTAAT CTTTTTTTCT TTTTCTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T

40 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAATC TTTTTTTTCT TTTTCTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

45

7. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,

daß das selektierbare Markergen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.

50 8. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,

daß das selektierbare Markergen ein Neomycinphosphotransferasegen ist.

55 9. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet,

daß das selektierbare Markergen strangabwärts der 5' NTR in die HCV-RNA integriert ist.

10. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 9, dadurch gekennzeichnet,

daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.

11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet,

daß es ein integriertes Reportergen aufweist.

12. HCV-RNA-Konstrukt nach Anspruch 11, dadurch gekennzeichnet,

daß das Reportergen ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasen), der GFP-Gene (green-fluorescence-protein-Gene), dem GUS-Gen (Glukuronidasen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen) ist.

13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 11, dadurch gekennzeichnet,

daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.

14. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 13, dadurch gekennzeichnet,

daß das Resistenzgen derart in das offene Leseraster der HCV-RNA inkloniert ist, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführbar ist.

15. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 14, dadurch gekennzeichnet,

daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionsprotein exprimieren.

16. Zellkultursystem nach einem der Ansprüche 1 bis 3 dadurch gekennzeichnet,

daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 15 ist.

17. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) hinterlegt sind.

18. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.

19. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

20. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung einer leberzellspezifischen Gentherapie.

21. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet,

daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

22. Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts gemäß einem der Ansprüche 4 bis 15, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, dadurch gekennzeichnet,

daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 15 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man

aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.

23. Verfahren nach Anspruch 22, dadurch gekennzeichnet,

5 daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 einschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Germaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

10

24. Verfahren zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV- Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, dadurch gekennzeichnet,

15

daß man mit einem Verfahren nach Anspruch 22 oder 23 eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt und isoliert,

daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt,

20

und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV- Vollängengenom oder ein HCV- Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

25. Zellkultur-adaptiertes HCV-RNA-Konstrukt mit hoher Replikationseffizienz, dadurch gekennzeichnet,

25

daß es durch Nukleotid- und/oder Aminosäure -Mutationen von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15 ableitbar ist und daß es mit einem Verfahren nach einem der Ansprüche 22 bis 24 erhältlich ist.

30

26. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25, dadurch gekennzeichnet,

daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist.

35

27. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25 oder 26, dadurch gekennzeichnet,

daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.

40

28. Zellkultur-adaptierte Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, dadurch gekennzeichnet,

45

daß sie mit einem Verfahren erhältlich ist, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt nach Anspruch 24 durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 15, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

50

29. Hepatitis C Viruspartikel oder virus-ähnliche Partikel dadurch gekennzeichnet,

daß sie mit einem Verfahren nach einem der Ansprüche 22-24 erhältlich sind.

55

30. Zellen, infiziert mit Hepatitis C Viruspartikeln oder virus-ähnliche Partikeln gemäß Anspruch 29.

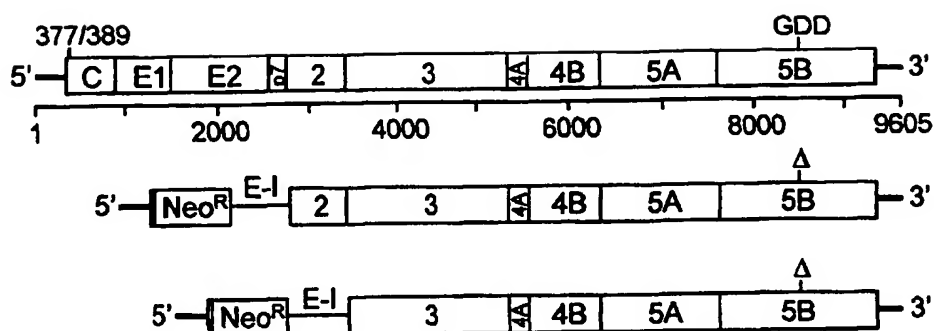
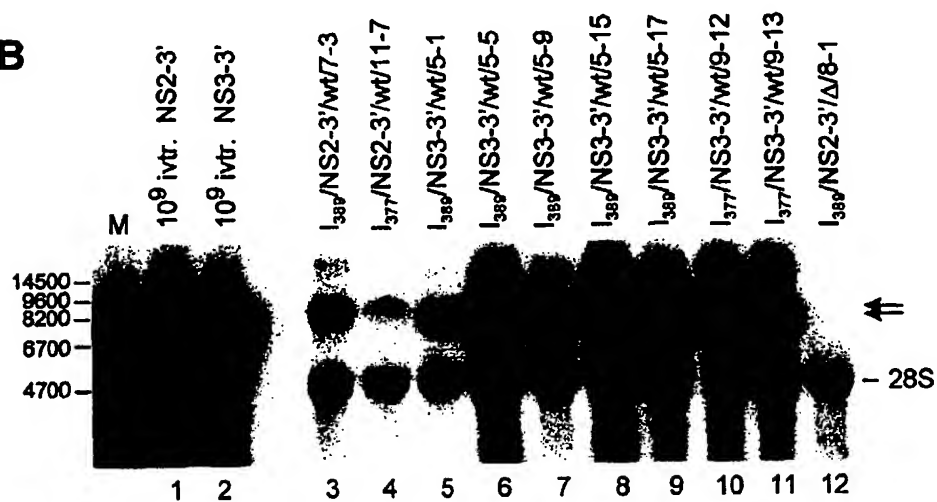
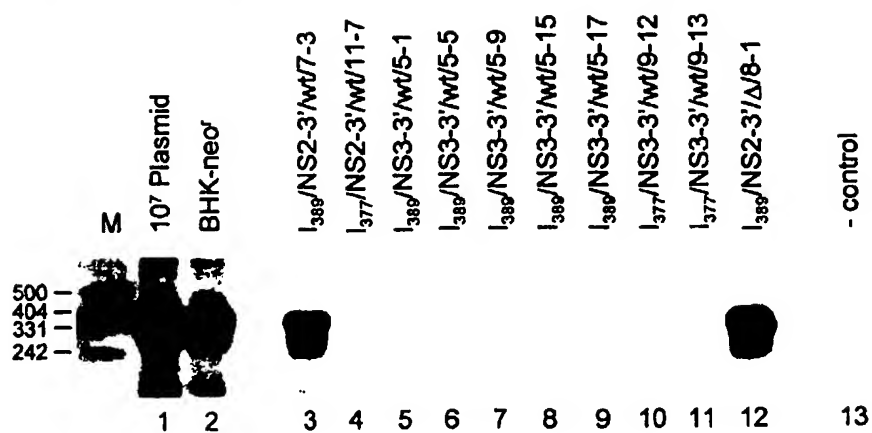
**A****B****C**

Fig. 1

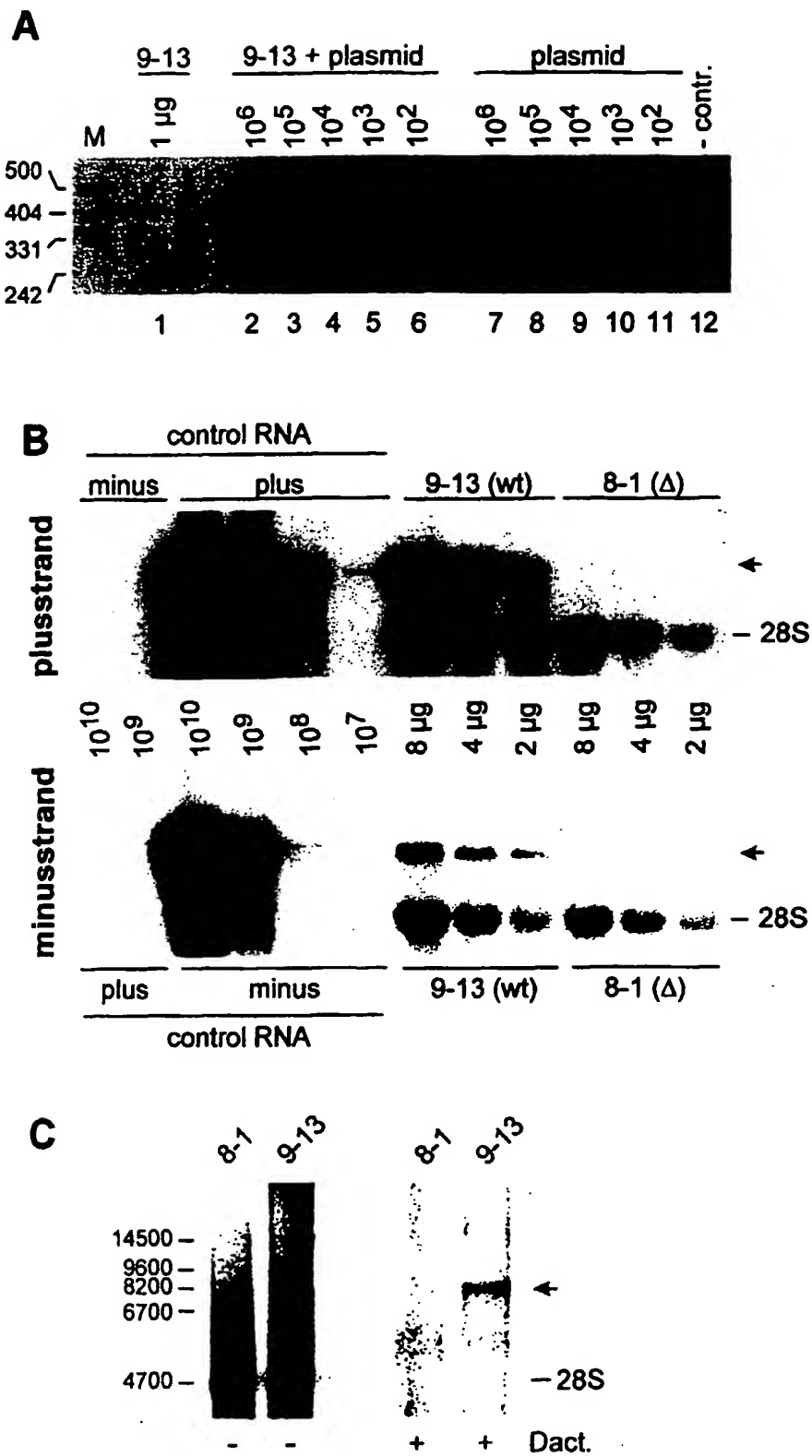


Fig. 2

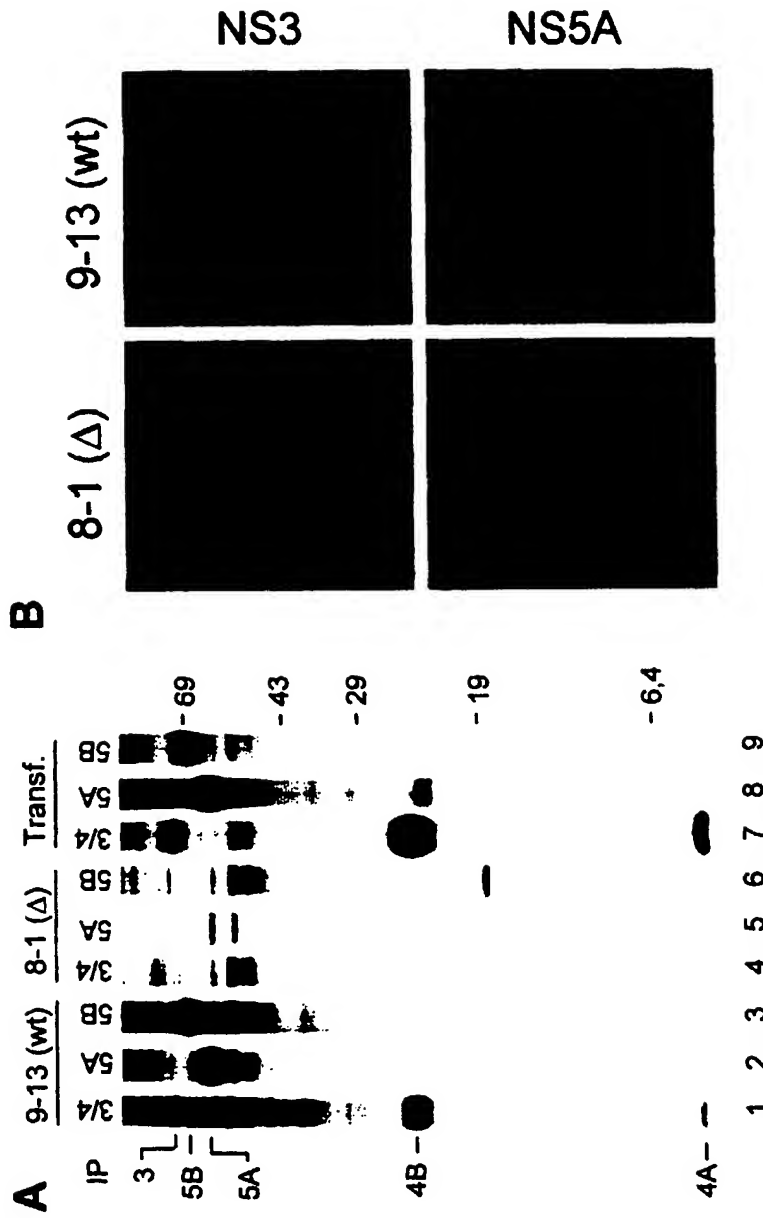


Fig. 3

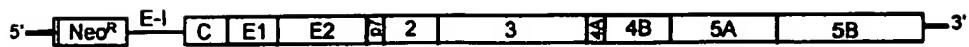


Fig. 4

A



B

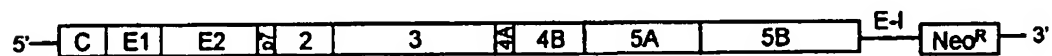


Fig. 5

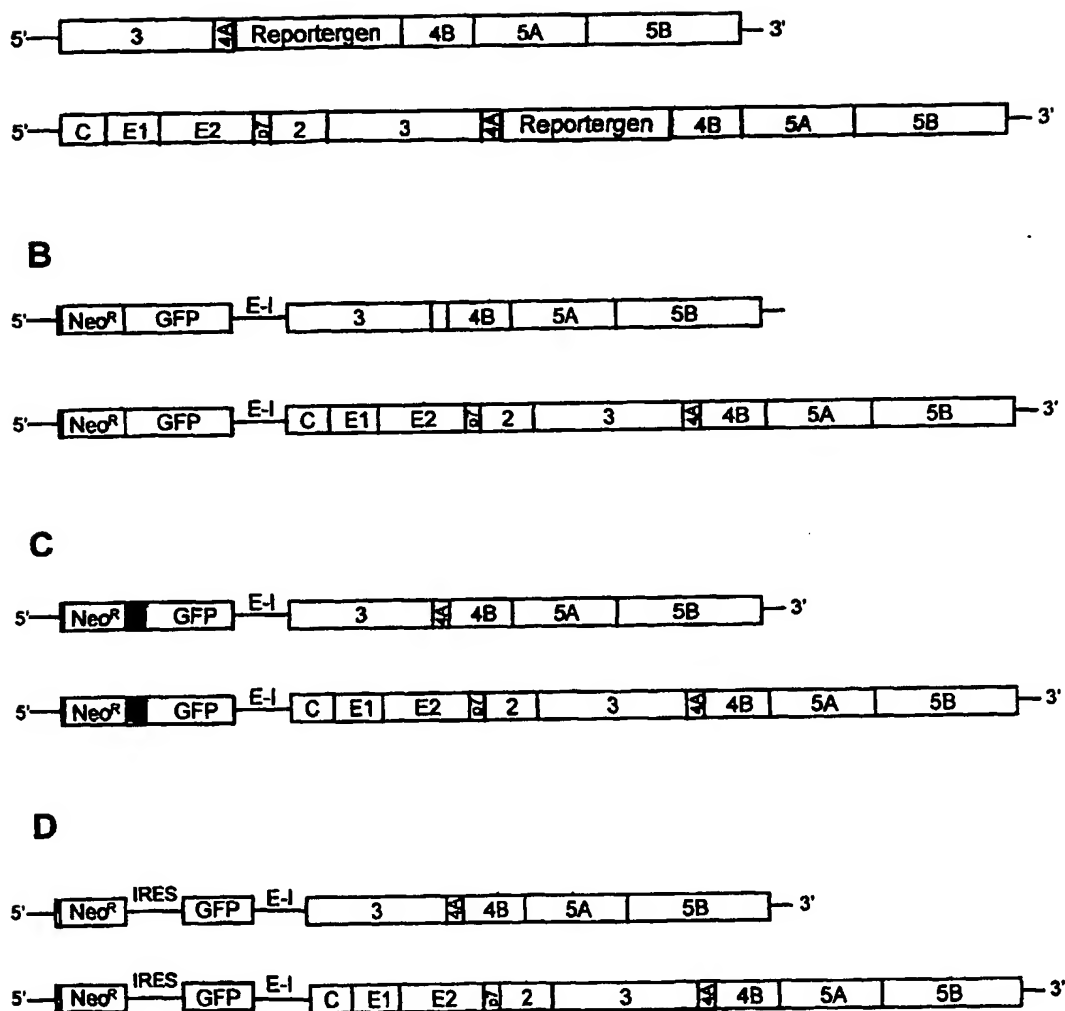


Fig. 6

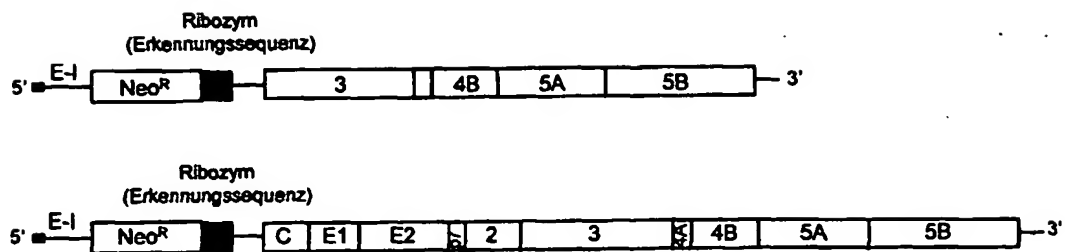


Fig. 7

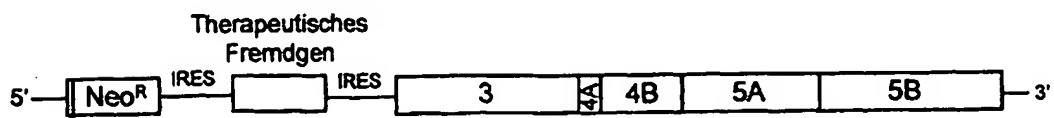


Fig. 8

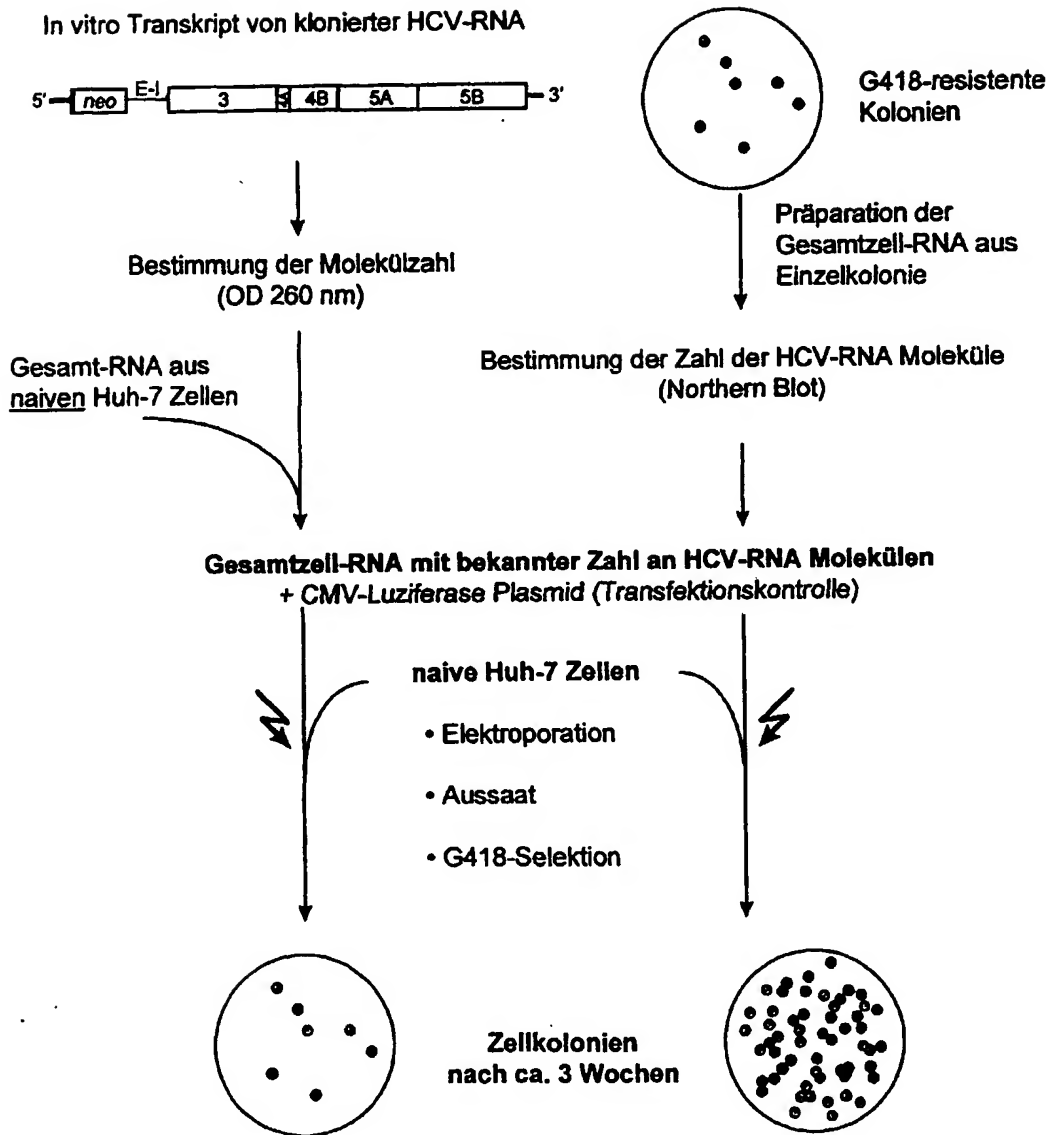


Fig. 9

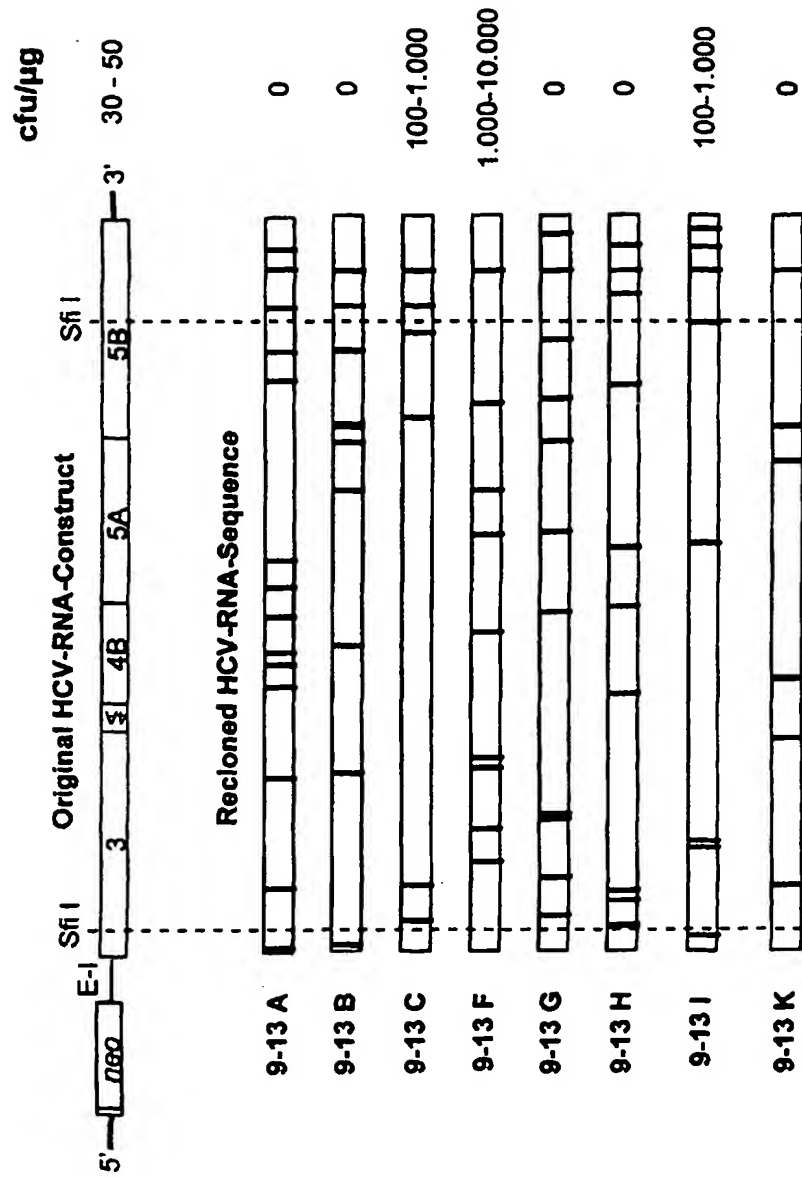


Fig. 10

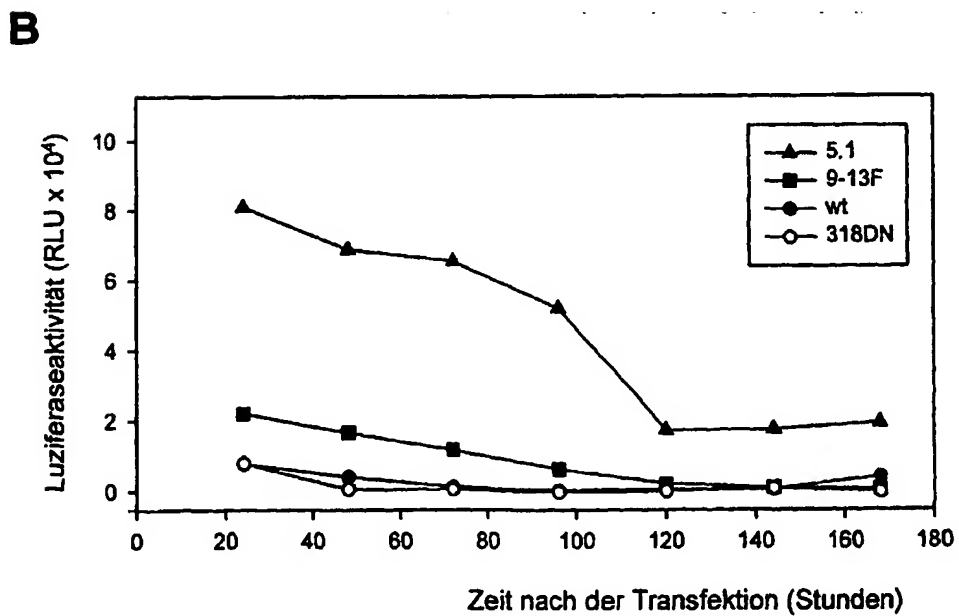
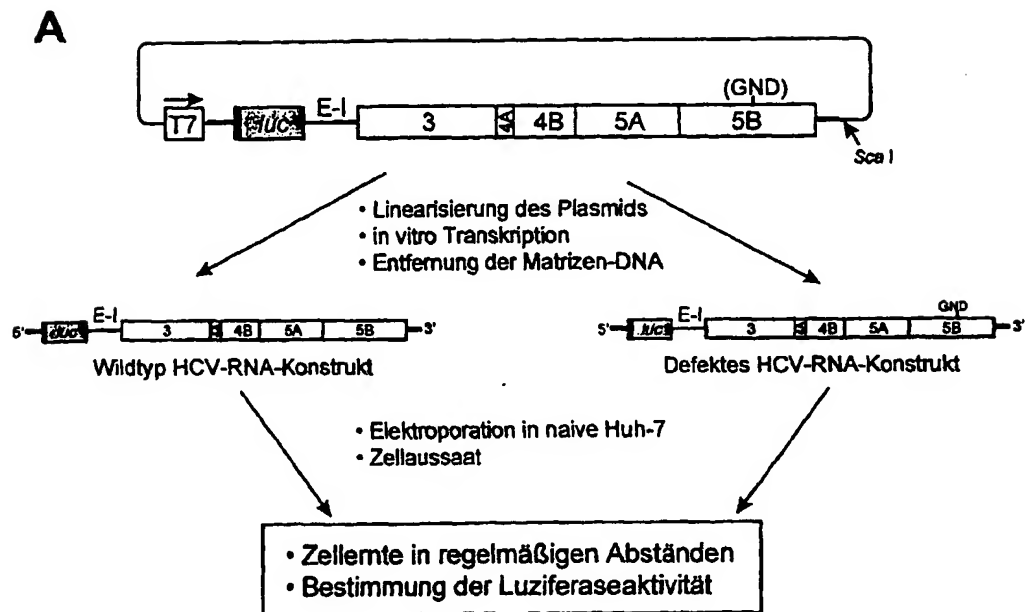
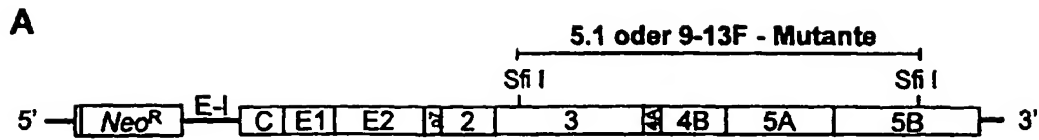


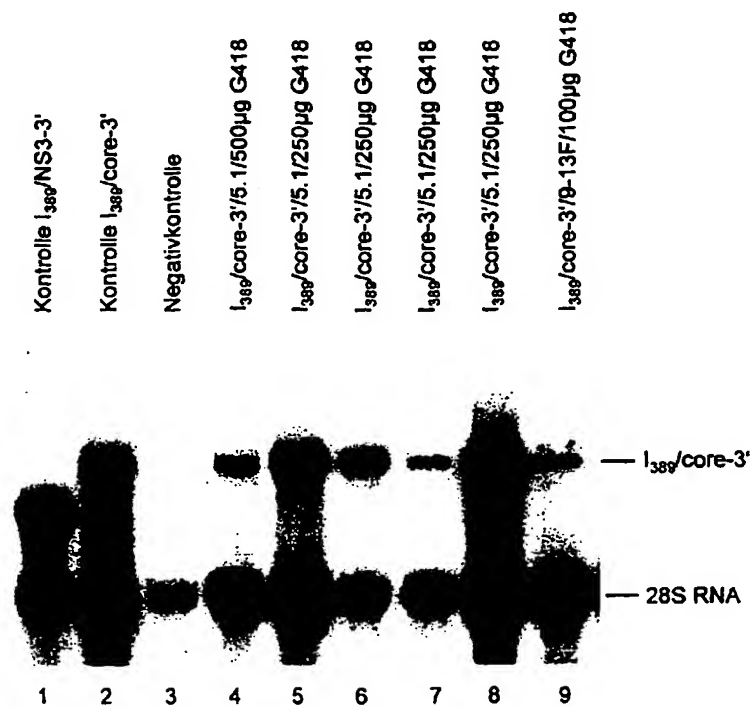
Fig. 11



**B**

G418-Konzentration:	500 µg/ml	250 µg/ml	100 µg/ml
Anzahl erhaltener Kolonien:	1	23	75

**C**



**Fig. 12**

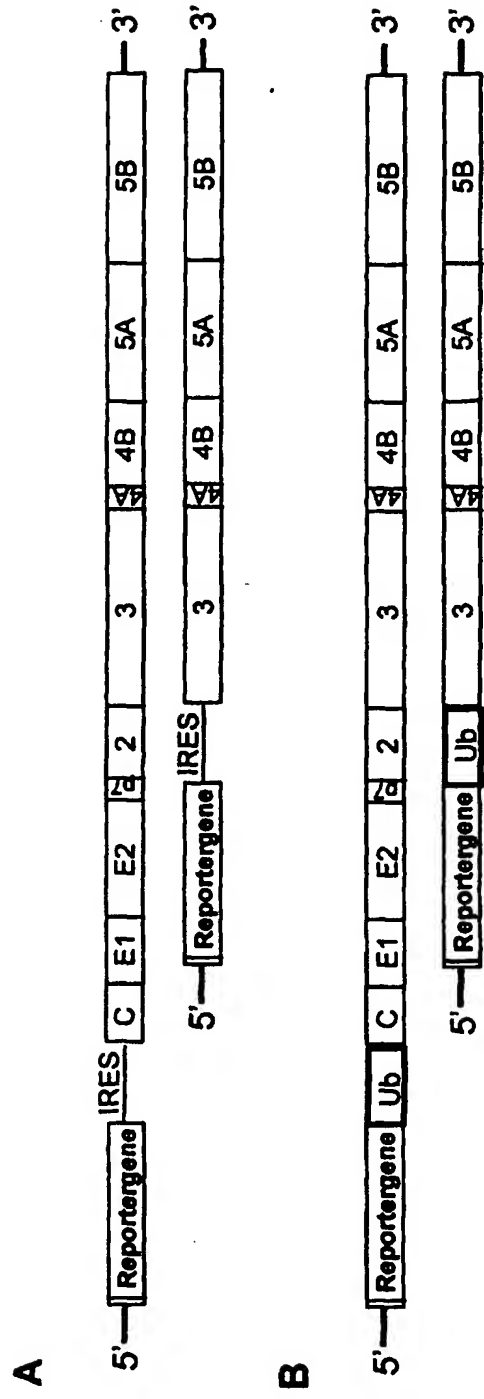


Fig. 13



Fig. 14

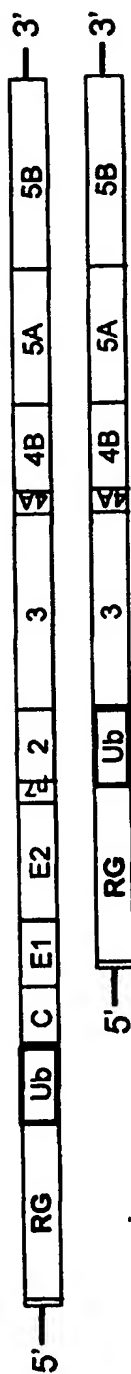


Fig. 15

(19) European Patent Office

(11) EP 1 043 399 A2

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of Publication:  
10/11/2000 Patent Gazette 2000/41

(51) Int. Cl.<sup>7</sup>: C12N 15/86, C12N 7/01,  
C12N 7/04, C12N 5/10,  
C07K 14/18, A61K 49/00,  
A61K 48/00

(21) Application Number: 00105929.4

(22) Date of Application: 3/23/2000

(84) Convention States Named:  
AT BE CH CY DE DK ES FI FR GB GR  
IE IT LI LU MC NL PT SE

(72) Inventor:  
**Bartenschlager, Ralf, Dr.**  
**55239 Gau-Odernheim**  
**(Germany)**

(30) Priority: 4/3/1999 DE 19915178

(74) Representative:  
**Rudolph, Ulrike, Dr.**  
**Patent Attorney**  
**In der Schanz 10**  
**69198 Schriesheim**  
**(Germany)**

(71) Applicant:  
**Bartenschlager, Ralf, Dr.**  
**55239 Gau-Odernheim (Germany)**

(54) **Hepatitis C Virus Cell Culture System**

(57) The hepatitis C virus (HCV) cell culture system pursuant to the invention consists of human hepatoma cells that are transfected with an HCV-RNA construct that comprises the HCV-specific RNA sections 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, and also at least one selectable marker gene (selection gene).

**THIS PAGE BLANK (USPTO)**

## Specification

[0001] This invention relates to a hepatitis C virus (HCV) cell culture system that comprises essentially eucaryotic cells that contain inserted HCV-specific genetic material, i.e. that are transfected with HCV-specific genetic material.

[0002] The hepatitis C virus (HCV) is one of the principal causes of chronic and sporadic liver diseases worldwide. Most HCV infections run their course with no recognizable clinical symptoms, but 80-90% of those infected are permanent carriers of the virus, and 50% of these permanent virus carriers have chronic liver inflammation of varying degrees of severity. About 20% of those chronically infected develop cirrhosis of the liver in the course of 10 to 20 years, on the basis of which primary hepatic cell carcinoma may develop. Chronic hepatitis C today is the chief indication for a liver transplant. As yet there is no causal therapy. The only therapy available at this time is administration of large doses of alpha interferon or a combination of alpha interferon and the purine nucleoside analog ribavirin. Of course only about 60% of all those treated respond to this therapy, and more than half of all these cases suffer renewed viremia after discontinuing treatment.

Because of its high prevalence, including in the industrialized nations, the severe consequences of chronic infections, and the lack of causal therapy, development of an HCV-specific chemotherapy is an important goal of pharmaceutical research and development. The main problem here so far is the lack of a suitable cell culture system that permits study of virus replication and pathogenesis in eucaryotic cells.

[0003] Because of the small amounts of virus in the blood and tissue, the lack of suitable cell culture systems or animal models (the chimpanzee is the only possible experimental animal so far), and the lack of efficient systems for producing virus-like particles, it has not yet been possible to study or clarify thoroughly the molecular composition of the HCV particle. The results available today can be summarized as follows: HCV is an enveloped plus-strand RNA virus with a particle diameter of 50-60 nm and an average density of 1.03-1.1 g/ml. It was cloned molecularly for the first time in 1989, and characterized (Choo et al., 1989: *Science*, 244, 359-362). HCV RNA has a length of about 9.6 kb (= 9600 nucleotides), positive polarity, and has a single open reading frame (ORF) that codes a linear polyprotein of about 3010 amino acids (see Rice 1996, in *Virology*, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), Vol. 1, pp. 931-960; Clarke 1997, *J. Gen. Virol.* 78, 2397; and Bartenschlager 1997, *Intervirology* 40, 378, and cf. Fig. 1 A). The polyprotein is cleaved into the mature and functionally active proteins during virus replication by cellular and viral proteases.

The proteins within the polyprotein are arranged as follows (from the amino end to the carboxyl end): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. The core protein is the main component of the nucleocapsid. The glycoproteins E1 and E2 are transmembrane proteins and main components of the viral envelope. They probably play an essential role in the fastening of the virus to the host cell. These three proteins core,

**THIS PAGE BLANK (USPTO)**

E1, and E2, make up the viral particle and are therefore called structural proteins. The function of the protein p7 is still unclear. The protein NS2 is probably the catalytic domain of the NS2-3 protease that is responsible for the processing between the proteins NS2 and NS3. The protein NS3 has two functions, namely protease activity in the amino-terminal domain that is essential for polyprotein processing, and an NTPase/helicase function in the carboxy-terminal domain that probably plays a role in the replication of the viral RNA. The protein NS4A is a cofactor of NS3 protease. The function of the protein NS4B is unknown.

[0004] The open reading frame is flanked at its 5' end by a non-translated region (NTR) about 340 nucleotides long that functions as an internal ribosome entry site (IRES), and at its 3' end by an NTR about 230 nucleotides long that most probably is important for genome replication. Such a 3' NTR is the object of Patent Application PCT/US 96/14033. The structural proteins in the amino-terminal quarter of the polyprotein are cleaved by the signal peptidase of the host cell. The non-structural proteins (NS) 2 to (NS) 5B are processed by two viral enzymes, namely by the NS2-3 and the NS3/4A proteinases. The NS3/4A proteinase is needed for all cleavages beyond the carboxy terminus of NS3. The role of NS4B is unknown. NS5A, a highly phosphorylated protein, seems to be responsible for the interferon resistance of various HCV genotypes (cf. Enomoto et al. 1995, *J. Clin. Invest.* 96, 224; Enomoto et al. 1996, *N. Engl. J. Med.* 334, 77; Gale Jr. et al. 1997, *Virology* 230, 217; Kaneko et al. 1994, *Biochem. Biophys. Res. Commun.* 205, 320; Reed et al., 1997, *J. Virol.* 71, 7187), and NS5B has been identified as RNA-dependent polymerase.

[0005] Using this information, initial diagnostic systems were developed that depend on either the detection of HCV-specific antibodies in the patient's serum or the detection of HCV-specific RNA by means of RT-PCR (= Reverse Transcription Polymerase Chain Reaction), and that are (have to be) used routinely and or as instructed for all blood stored in the meantime.

[0006] Since the first description of the genome in 1989, numerous partial and complete sequences of HCV have been cloned and characterized using the PCR method. Comparison of these sequences shows high variability of the viral genome, particularly in the region of the NS5B gene, which has ultimately led to classification into 6 genotypes, which themselves have also been subclassified again into subtypes a, b, and c. The genomic variation is not distributed uniformly over the entire genome. Thus, the 5'NTR and parts of the 3'NTR are highly preserved, while certain coding sequences sometimes vary very sharply, especially the envelope proteins E1 and E2.

[0007] The cloned and characterized partial and complete sequences of the HCV genome have also been studied with regard to suitable points of attack for a prospective antiviral therapeutic agent. Three viral enzymes have been found here that are available as such points of attack. These are (1) the NS3/4A protease complex, (2) the NS3 helicase, and (3) the NS5B RNA-dependent RNA polymerase. The NS3/4A protease complex and the NS3 helicase have already been crystallized and their three-dimensional structures have been clarified (Kim et al. 1996, *Cell*, 87, 343; Yem et al., 1998, *Protein Science*, 7, 837;



**THIS PAGE BLANK (USPTO)**

Love et al., 1996, *Cell*, 87, 311; Kim et al., 1998, *Structure*, 6, 89; Yao et al., 1997, *Nature Structural Biology*, 4, 463; Cho et al., 1998, *J. Biol. Chem.*, 273, 15045); this has not yet been accomplished for the NS5B RNA-dependent RNA polymerase.

Although significant points of attack are defined with these enzymes for the development of a therapeutic treatment of chronic HCV infection, and although suitable inhibitors are being sought intensively around the world using both 'rational drug design' and 'high throughput screens', therapy development suffers from a large deficit, namely the lack of cell culture systems or simple animal models that would permit the reliable detection by simple laboratory methods of HCV RNA or HCV antigens. The lack of such cell culture systems is also the fundamental reason why the understanding of HCV replication up to the present time is still very sketchy, and in broad areas only hypothetical.

[0008] Although in the opinion of well-informed specialists there exists a close evolutionary relationship between HCV and the flaviviruses and pestiviruses, and autonomously replicating RNAs are described for these that can be brought directly into various cell lines for replication and then show relatively high yields (see Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998 *J. Virol.* 72, 5318), similar attempts with HCV have so far been unsuccessful.

[0009] Actually it is known from various publications that cell lines or primary cell cultures can be infected with high-titer patient's serum containing HCV (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences*, USA, 90, 6037-6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219-7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 and citations therein; Ito et al. 1996, *Journal of General Virology*, 77, 1043-1054), these virus-infected cell lines or cell cultures nevertheless do not permit the direct detection of HCV RNA or HCV antigens. Neither is the viral RNA in these cells detectable in a Northern Blot (a standard procedure for the quantitative determination of RNA) nor are the viral proteins detectable in a Western Blot or by immunoprecipitation. It has been at all possible only by very tedious and indirect methods to detect HCV replication. These unfavorable circumstances show clearly that replication in these known virus-infected cell lines or cell cultures is absolutely insufficient.

[0010] Furthermore it is known from the publications by Yoo et al. (1995, *Journal of Virology*, 69, 32-38) and by Dash et al., (1997, *American Journal of Pathology*, 151, 363-373) that hepatoma cell lines can be transfected with synthetic HCV RNA that was obtained by in vitro transcription of cloned HCV genome. In both publications the authors assume the fundamental concept that the viral HCV genome is a plus strand RNA that functions directly as mRNA after insertion into the cell, to which ribosomes are attached and form virus proteins in the course of translation processes, from which ultimately new HCV particles (can) form. This viral replication, i.e. these newly formed HCV viruses and their RNA, were detected by means of RT PCR. The published results of the RT PCR carried out, however, are evidence that the efficiency of HCV replication in the described HCV-transfected hepatoma cells is only very slight and in any case is not sufficient to measure even qualitatively, not to mention quantitatively, variations in the

**THIS PAGE BLANK (USPTO).**

rate of replication after the selective action of prospective antiviral therapeutic agents. Meanwhile it is also disclosed in the state of the art (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999) that highly-preserved 3' NTR is essential for viral replication, which is in clear contradiction to the assertions of Yoo et al. and Dash et al., who used exclusively HCV genome with shortened 3' NTRs for their experiments, not knowing of the authentic 3' end of the HCV genome.

[0011] The purpose of this invention is to make available an HCV cell culture system in which the viral RNA in the transfected cells replicates autonomously and with such high efficiency that variations in the rate of replication after the selective action of virus-specific and particularly HCV-specific prospective antiviral therapeutic agents can be measured qualitatively and quantitatively and using current measurement procedures customary in the laboratory.

[0012] A solution of this problem consists of making available a cell culture system of the type mentioned initially in which the eucaryotic cells are human cells, particularly hepatoma cells, that are derived preferably from a commercial hepatoma cell line, but may also be obtained from a corresponding primary cell culture, and in which the inserted HCV-specific genetic material is an HCV-RNA construct that comprises essentially the HCV-specific RNA sections 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, preferably in the order given, and also comprises at least one selectable marker gene (selection gene). "NTR" in this context and below means "non-translated region" and is well known to one skilled in the art as a concept and as an abbreviation. The phrase "HCV-RNA construct" in this context and below comprises both constructs that contain the complete HCV genome and those that contain only a portion thereof, i.e. an HCV subgenome.

A preferred variant of the cell culture system pursuant to the invention that has proved itself well in practice is deposited with the DSMZ, German Collection of Microorganisms and Cell Cultures GmbH in Braunschweig, Germany, under Number DSM ACC2394 (laboratory designation HuBl 9-13).

[0013] The cell culture system pursuant to the invention already makes available for the first time an in vitro system in which HCV RNA replicates autonomously and is expressed in sufficiently large amounts for quantitative determination of the amounts both of HCV RNA and of the HCV-specific proteins to be able to be carried out by conventional and reliably accurate biochemical measurement methods. This means that for the first time an approximately authentic cell-based HCV replication system is available that is urgently needed for the development and testing of antiviral pharmaceuticals. This test system now provides the ability to identify potential sites of attack for effective HCV-specific therapy and to develop and evaluate HCV-specific chemotherapeutic agents.

[0014] The invention is based on the surprising determination that efficient replication of HCV RNA takes place in cells only when transfection is performed with an HCV RNA construct that comprises at least the 5' and the 3' non-translated regions (NTR) and the

**THIS PAGE BLANK (USPTO)**

non-structural proteins (NS) 3 to 5B, and additionally has a selectable marker gene (selection gene). Apparently the structural genes are not of major importance for the occurrence of replication, while on the other hand efficient replication of HCV RNA apparently takes place only when the transfected cells are subjected to constant selection pressure that is mediated by the selectable marker gene (selection gene) associated with the HCV RNA. The marker gene (selection gene) thus appears on the one hand to provoke the selection of the cells in which the HCV RNA replicates productively, and on the other hand it appears to increase substantially the efficiency of RNA replication.

[0015] Also an object of the invention is a cell-free HCV RNA construct that is distinguished by the fact that it comprises the HCV-specific RNA sections 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, preferably in the order given, and also comprises a selectable marker gene (selection gene).

[0016] The expressions 5' NTR and NS3 and NS4A and NS4B and NS5A and NS5B and 3' NTR in this context include any nucleotide sequence that is described in the state of the art as a nucleotide sequence for the particular functional section of the HCV genome in question.

[0017] Providing such an HCV RNA construct for the first time makes possible detailed analysis of HCV replication, pathogenesis, and evolution in cell cultures. HCV-specific viral RNA can be produced selectively in any amount - as the complete genome or as a subgenome - and the ability exists to manipulate the RNA construct and with it to investigate and to clarify the HCV functions on the gene level.

[0018] Since all of the HCV enzymes studied so far as principal points of attack for therapy, namely NS3/4A protease, NS3 helicase, and NS5B polymerase, are contained in the HCV RNA construct pursuant to the invention, it can be used for all corresponding studies.

[0019] An embodiment of the HCV RNA construct that proved itself well in practical use is distinguished by the fact that it comprises the nucleotide sequence according to sequence protocol SEQ ID NO:1. Other variants of embodiment with comparably favorable characteristics for use in practice are characterized by the fact that they comprise a nucleotide sequence according to sequence protocol SEQ ID NO:2 or SEQ NO:3 or SEQ ID NO:4 or SEQ ID NO:5 or SEQ ID NO:6 or SEQ ID NO:7 or SEQ ID NO:8 or SEQ ID NO:9 or SEQ ID NO:10 or SEQ ID NO:11.

[0020] There is the ability to provide the HCV subgenome construct pursuant to the invention with a 3' NTR that has a nucleotide sequence so far unknown for this purpose in the state of the art, namely a nucleotide sequence that is selected from the group of nucleotide sequences (a) to (i) listed below:

**THIS PAGE BLANK (USPTO)**

- (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA  
GCTTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
AGACTGGAGT
- (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCAGACT GGAGC
- (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCAGACT GCAGC
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCAG ACTGGAGT
- (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCAG ACTGGAGT
- (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TTGAGAGAGA GAGTCTCACT  
CTGTTGCCCA GACTGGAGT
- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA  
GTCTCACTCT GTTGCCAGA CTGGAGT
- (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTTAAT CTTTTTTTTT TTTTCCTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTTAATC TTTTTTTTTT TTTCTTTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

**THIS PAGE BLANK (USPTO)**

The selectable marker gene (selection gene) contained in the HCV RNA constructs pursuant to the invention is preferably a resistance gene, especially an antibiotic resistance gene.

This has the benefit that cells transfected with this construct can easily be selected from the untransfected cells, for example in the case of an antibiotic resistance gene, by adding the particular antibiotic to the cell culture medium. 'Antibiotic' in this context means any substance that prevents the non-transfected host cells or the cells in which HCV RNA replicates with only low efficiency from living or growing, particularly cell poisons such as puromycin, hygromycin, zeocin, bleomycin, or blasticidin.

[0021] A preferred selectable marker gene (selection gene) or resistance gene that has proved itself well in practice is the neomycin phosphotransferase gene.

[0022] An alternative to antibiotic resistance genes, for example, is the thymidinekinase gene, with which an HAT selection can be carried out.

[0023] The position of the selectable marker gene (selection gene) or of the preferred resistance gene or the particularly preferred antibiotic resistance gene in the HCV RNA construct is preferably beyond the HCV 5' NTR, i.e. downstream from the 5' NTR and upstream from the HCV reading frame. However, insertion in the region of the 3' NTR or at another point in the HCV genome or subgenome, for example within the polyprotein, is also conceivable.

[0024] In an alternative embodiment of the HCV RNA construct pursuant to the invention, the selectable marker gene (selection gene), particularly an antibiotic resistance gene, is connected to the HCV RNA or HCV genome or subgenome sequence through a ribozyme or a recognition site for a ribozyme.

[0025] This provides the advantage that after selection of the cells in which HCV RNA is productively replicated, the resistance gene in the cell clones obtained can be removed by ribozyme-mediated cleavage of the HCV subgenome sequence, namely by activation of the cloned ribozyme, or in case of a construct with a recognition site for a ribozyme, by insertion of the ribozyme into the cells (for example by transfection of a ribozyme construct or infection with a viral expression vector in which the appropriate ribozyme was used). In this way an authentic HCV genome construct is obtained with no resistance gene, which is capable of forming authentic infectious virus particles.

[0026] Another preferred embodiment of the HCV RNA construct pursuant to the invention is characterized by the fact that the construct has at least one integrated reporter gene.

[0027] A reporter gene in the following means any gene whose presence can be detected readily after transfer into a target organism, and in general by simple biochemical or histochemical methods, i.e. that codes for a protein that can be detected and quantified

**THIS PAGE BLANK (USPT**

readily and reliably even in small amounts by customary laboratory measurement methods.

[0028] This variant of the HCV RNA construct has the advantage that the extent of replication of this construct can be measured easily and quickly by conventional laboratory methods by reference to the reporter gene product.

[0029] The reporter gene is preferably a gene from the group of luciferase genes, the CATG gene (chloramphenicol acetyltransferase gene), the lacZ gene (beta-galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene), or the SEAP gene (secreted alkaline phosphatase gene). These reporter genes or their products, namely the corresponding reporter proteins, can be determined, for example, by fluorescence, chemiluminescence, colorimetrically, or by immunological methods (e.g. ELISA).

[0030] However, a surrogate marker gene is also practical as the reporter gene. Among these in this context are genes that code for cellular proteins, nucleic acids, or - generally - for functions that are subject to a variation dependent on the virus replication, and that consequently are either repressed or activated in cells in which the HCV or the HCV RNA construct multiplies. In other words, the reduction or activation of this function is a replacement marker for virus replication or replication of the HCV RNA construct.

[0031] The positions of reporter gene and selectable marker gene (selection gene) can be chosen so that a fusion protein of the two gene products formed is expressed. In this case the advantageous ability exists to arrange these two genes in the HCV RNA construct so that their two expressed proteins are first fused through a cutting site for a protease (e.g. ubiquitin) or through a self-cleaving peptide (e.g. the 2A protein of picornaviruses), and only later are separated again proteolytically.

However, these two positions can just as well be separated from one another in such a way that the two genetic products are expressed separately. (For example in the order: Marker or resistance gene - Internal ribosome binding site - Reporter gene).

In the case of the reporter gene, a variant of embodiment has proved itself particularly well in which the reporter gene is cloned into the open reading frame of the HCV genome or subgenome specifically in such a way that it is converted to an active form only after proteolytic processing.

[0032] The cell culture system pursuant to the invention in all its variations can be used for diverse purposes. These include:

- The discovery of antivirally active substances. These can be, for example: organic compounds that intervene directly or indirectly in virus multiplication (for example, inhibitors of viral proteases, of NS3 helicase, of NS5B RNA-dependent RNA polymerase), antisense oligonucleotides that hybridize at any target

**THIS PAGE BLANK** (USPTO)

sequence within the HCV RNA construct (e.g. the 5' NTR) and lead directly or indirectly to an influence on virus multiplication, for example by reduction of translation of the HCV polyprotein, or ribozymes that cleave any HCV RNA sequence and thus impair virus multiplication.

- The evaluation of any type of antivirally active substances in cell culture. Such substances, for example, can be found by 'rational drug design' or 'high-throughput screening' on the isolated purified enzyme. Evaluation means especially determination of the inhibiting properties of the corresponding substance and its mechanism of action.
- The identification of new targets of attack, of viral or cellular origin, for HCV-specific antiviral therapy. For example, if a cellular protein is essential for viral replication, viral replication can likewise be controlled by inhibiting this cellular protein. The discovery of such auxiliary factors is likewise possible with the system pursuant to the invention.
- Use for resistance determination. It can be assumed that resistance to therapy occurs because of the high mutation rate of the HCV genome. Such resistance, which is of great importance in the clinical licensing of a substance, can be determined with the cell culture system pursuant to the invention. Cell lines in which the HCV RNA construct or the HCV genome or subgenome replicates are incubated with increasing concentrations of the appropriate substance, and the replication of the viral RNA is determined either by means of an introduced reporter or by qualitative or quantitative determination of the viral nucleic acids or proteins. Resistance exists when no inhibition of replication is observed with normal drug concentration. The nucleotide or amino acid replacements responsible for resistance to therapy can be determined by recloning of the HCV RNA (e.g. by RT-PCR) and sequence analysis. Their causality for resistance to therapy can be proven by cloning the corresponding replacement(s) into the original construct.
- The production of authentic viral proteins (antigens) for the development and/or evaluation of diagnostic products. The cell culture system pursuant to the invention also permits the expression of HCV antigens in cell cultures. These antigens in principle can also be used for the design of diagnostic detection methods.
- The production of HCV viruses and virus-like particles, particularly for the development or production of therapeutic agents and vaccines as well as for diagnostic purposes. In particular, cell culture-adapted complete HCV genomes that can be produced with the cell culture system pursuant to the invention are able to replicate in cell cultures with high efficiency. These genomes have all of the functions of HCV and are therefore able to produce infectious viruses.

**THIS PAGE BLANK (USPTO)**

[0033] The HCV RNA construct pursuant to the invention can also be used by itself in all of its variations for diverse purposes. Among them in particular are:

- The construction of attenuated hepatitis C viruses or HCV-like particles and their production in cell cultures; attenuated HCV or HCV-like particles can be produced by random mutations, for example point mutations, deletions or insertions, i.e., viruses or virus-like particles with full replication competence but reduced or absent pathogenicity. Such attenuated HCV or HCV-like particles are especially useful as vaccines.
- The construction of HCV RNA constructs with integrated foreign genes, for example for use as liver cell-specific gene ferries in gene therapy. Because of the pronounced liver cell tropism of HCV and the ability to replace parts of the genome with heterologous sequences, HCV RNA constructs can be prepared in which the structural proteins are replaced by a therapeutically active gene, for example. The HCV RNA construct thus obtained is inserted into cells, preferably by transfection, that constitutionally or inducibly express the missing HCV functions, for example the structural proteins. By this technique, known to one skilled in the art by the term 'transcomplementation', virus particles can be produced in which the HCV RNA construct is incorporated. The particles thus obtained can be used for the preferential infection of liver cells. The therapeutically active foreign gene in them is expressed and then deploys its therapeutic action.
- The discovery of permissive cells, i.e. cells in which productive virus multiplication occurs. For this purpose either one of the aforementioned HCV RNA genome constructs that has the ability to form complete infectious viruses is used, or one of the aforementioned HCV subgenome constructs is used that of course is first transfected into a cell line according to the aforementioned example that expresses the missing functions constitutionally or inducibly. In all these cases virus particles are formed that also carry a resistance gene and/or a reporter gene, in addition to the HCV sequence. To find cells in which HCV can replicate, these cells are infected with the viruses thus prepared and are subjected to antibiotic selection, or depending on the HCV RNA construct, they are studied by detecting the expression of the reporter gene. Since antibiotic resistance or expression of the reporter gene can be detected only when the HCV RNA construct replicates, the cells found in this way must be permissive. Almost any cell lines or primary cell cultures can be tested and permissivity can be revealed in this way.

[0034] The cell culture system pursuant to the invention also permits the discovery of HCV RNA constructs in which there is an increase of replication efficiency because of mutations, which occur either randomly in the course of HCV RNA replication or are introduced selectively into the construct. Such mutations that lead to a change of replication of the HCV RNA construct are known to one skilled in the art as adaptive mutations. The invention therefore comprises also methods for obtaining cell culture-

**THIS PAGE BLANK (USPTO)**

adapted mutants of an HCV RNA construct pursuant to the invention according to the above description, with the mutants having higher replication efficiency than the original HCV RNA construct. It also comprises a method for producing mutants of a full-length HCV RNA genome or of a partial HCV RNA genome or an HCV RNA construct with increased replication efficiency, as well as cell culture-adapted mutants of HCV RNA constructs, full-length HCV genomes, and partial HCV genomes with higher replication efficiency than the original constructs, or partial or full-length genomes.

[0035] The method pursuant to the invention for obtaining cell culture-adapted mutants of an HCV RNA construct pursuant to the invention, with the mutants having higher replication efficiency than the HCV RNA construct, is characterized by the fact that a cell culture system pursuant to Claim 1 in which the inserted HCV-specific genetic material is an HCV RNA construct with selection gene pursuant to one of the claims 4 to 19, is cultivated in the selection medium corresponding to the selection gene, the grown cell line is harvested, and the HCV RNA constructs are isolated from these cell clones.

[0036] In an advantageous refinement of this production method, the isolated HCV RNA constructs are again passed at least once more, namely inserted in cells of a cell culture system pursuant to Claim 1, the cell culture system pursuant to Claim 1 that is obtained, in which the inserted HCV-specific genetic material is the isolated HCV RNA construct with selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the grown cell clone is harvested, and the HCV RNA constructs are isolated therefrom.

The degree of adaptive mutations and with it the degree of replication efficiency in the HCV RNA constructs in question can be further increased by this variant of the method.

[0037] The method pursuant to the invention for producing mutants of a full-length HCV genome or of a partial HCV genome or of any HCV RNA construct with higher replication efficiency than the original full-length HCV genome or partial genome or HCV RNA construct is distinguished by the fact that a cell culture-adapted mutant of an HCV RNA construct is prepared using one of the two production methods mentioned above, it is isolated from the cells, cloned by methods known in the state of the art, and sequenced, and the type, number, and positions of the mutations are determined by comparison with the nucleotide and amino acid sequences of the original HCV RNA construct, and these mutations are then introduced into an (isolated) full-length HCV genome or partial genome or any HCV RNA construct either by selective mutagenesis or by replacement of sequence sections that the mutations in question contain.

To detect or verify the mutations that actually bring about a change of replication and in particular an increase of replication, a test can be carried out in which the intended nucleotide and/or amino acid replacement is introduced into the original HCV RNA construct and this in turn is inserted in cell culture. If the inserted mutation actually leads to an increase of replication, in the case of an HCV RNA construct with selectable marker gene the number of resistant cell clones in the artificially mutated construct should be clearly higher than in the untreated construct. In the case of a construct with a

**THIS PAGE BLANK (USPTO)**

reporter gene, the activity and the amount of the reporter should be clearly higher in the artificially mutated construct than in the untreated one.

[0038] The cell culture-adapted HCV RNA constructs pursuant to the invention with high replication efficiency are characterized by the fact that they can be derived from an HCV RNA construct pursuant to one of the claims 4 to 19 by nucleotide and/or amino acid exchanges and that they are available by one of the two methods of production mentioned above.

[0039] These cell culture-adapted HCV RNA constructs can be used to prepare any HCV RNA constructs or full-length or partial HCV genomes with increased replication efficiency. Either constructs with a selectable resistance gene or constructs with no such gene or with a non-selectable reporter gene (e.g. luciferase) can be prepared, since because of the very high replication efficiency of the cell culture-adapted HCV RNA construct, its replication can be demonstrated even in non-selected cells.

The cell culture-adapted mutants of an HCV RNA construct or of a full-length HCV genome or of a partial HCV genome pursuant to the invention with higher replication efficiency than the original HCV RNA construct or than the original full-length HCV genome are characterized by the fact that they are available by a method in which the type and number of mutations are determined in a cell culture-adapted HCV RNA construct by sequence analysis and sequence comparison, and these mutations are introduced into an HCV RNA construct, especially into an HCV RNA construct pursuant to one of the claims 4 to 19, or into an (isolated) full-length HCV RNA genome, either by selective mutagenesis or by replacement of sequence sections that contain the mutations in question.

[0040] A group of especially preferred HCV RNA constructs, full-length HCV genomes, and partial HCV genomes with high and very high replication efficiency and consequently very good suitability for practical use is characterized by the fact that it has one or more or all of the amino acid or nucleotide replacements listed in Table 3 and/or one or more of the following amino acid replacements: 1283 arg -> gly, 1383 glu -> ala, 1577 lys -> arg, 1609 lys -> glu, 1936 pro -> ser, 2163 glu -> gly, 2330 lys -> glu, 2442 ile -> val. (The numbers refer to the amino acid positions of the polyprotein of the HCV isolate con1; see Table 1).

**Special characteristics of the sequences indicated in the sequence protocols:**

**SEQ ID NO: 1**

[0041]

Name: I389/Core-3'/wt

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region

**THIS PAGE BLANK (USPTO)**

2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-10842: HCV polyprotein from core to non-structural protein 5B
5. 1813-2385: HCV core protein; structural protein
6. 2386-2961: Envelope protein 1; structural protein
7. 2962-4050: Envelope protein 2; structural protein
8. 4051-4239: Protein p7
9. 4240-4890: Non-structural protein 2 (NS2); HCV NS2-3 protease
10. 4891-6783: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
11. 6784-6945: Non-structural protein 4A (NS4A); NS3 protease cofactor
12. 6946-7728: Non-structural protein 4B (NS4B)
13. 7729-9069: Non-structural protein 5A (NS5A)
14. 9070-10842: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
15. 10846-11076: HCV-3' non-translated region

SEQ ID NO: 2

[0042]

Name: I337/NS2-3'/wt  
Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1181: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1190-1800: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1801-8403: HCV polyprotein from non-structural protein 2 to non-structural protein 5B
5. 1801-2451: Non-structural protein 2 (NS2); HCV NS2-3 protease
6. 2452-4344: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
7. 4345-4506: Non-structural protein 4A (NS4A); NS3 protease cofactor
8. 4507-5289: Non-structural protein 4B (NS4B)
9. 5290-6630: Non-structural protein 5A (NS5A)
10. 6631-8403: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
11. 8407-8637: HCV 3' non-translated region

**THIS PAGE BLANK (USPTO)**

**SEQ ID NO: 3**

[0043]

Name: I389/NS3-3'/wt

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-7767: HCV polyprotein from non-structural protein 3 to non-structural protein 5B
5. 1813-3708: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
6. 3709-3870: Non-structural protein 4A (NS4A); NS3 protease cofactor
7. 3871-4653: Non-structural protein 4B (NS4B)
8. 4654-5994: Non-structural protein 5A (NS5A)
9. 5995-7767: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
10. 7771-8001: HCV 3' non-translated region

**SEQ ID NO: 4**

[0044]

Name: I337/NS3-3'/wt

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1181: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1190-1800: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1801-7758: HCV polyprotein from non-structural protein 3 to non-structural protein 5B
5. 1801-3696: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
6. 3697-3858: Non-structural protein 4A (NS4A); NS3 protease cofactor
7. 3859-4641: Non-structural protein 4B (NS4B)
8. 4642-5982: Non-structural protein 5A (NS5A)
9. 5983-7755: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
10. 7759-7989: HCV 3' non-translated region



**THIS PAGE BLANK** (USPTO,

## SEQ ID NO: 5

[0045]

Name: I389/NS2-3'/wt

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-8418: HCV polyprotein from non-structural protein 2 to non-structural protein 5B
5. 1813-2463: Non-structural protein 2 (NS2); HCV NS2-3 protease
6. 2464-4356: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
7. 4357-4518: Non-structural protein 4A (NS4A); NS3 protease cofactor
8. 4519-5301: Non-structural protein 4B (NS4B)
9. 5302-6642: Non-structural protein 5A (NS5A)
10. 6643-8415: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
11. 8419-8649: HCV 3' non-translated region

## SEQ ID NO: 6

[0046]

Name: I389/NS3-3'/9-13F

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-7767: HCV polyprotein from non-structural protein 3 to non-structural protein 5B of the cell culture-adapted mutant 9-13F
5. 1813-3708: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
6. 3709-3870: Non-structural protein 4A (NS4A); NS3 protease cofactor
7. 3871-4653: Non-structural protein 4B (NS4B)
8. 4654-5994: Non-structural protein 5A (NS5A)
9. 5995-7767: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase

**THIS PAGE BLANK (US**

7771-8001: HCV 3' non-translated region

**SEQ ID NO: 7**

**[0047]**

Name: I389/core-3'/9-13F

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-10842: HCV polyprotein from core to non-structural protein 5B of the cell culture-adapted mutant 9-13F
5. 1813-2385: HCV core protein; structural protein
6. 2386-2961: Envelope protein 1; structural protein
7. 2962-4050: Envelope protein 2; structural protein
8. 4051-4239: Protein p7
9. 4240-4890: Non-structural protein 2 (NS2); HCV NS2-3 protease
10. 4891-6783: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
11. 6784-6945: Non-structural protein 4A (NS4A); NS3 protease cofactor
12. 6946-7728: Non-structural protein 4B (NS4B)
13. 7729-9069: Non-structural protein 5A (NS5A)
14. 9070-10842: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
15. 10846-11076: HCV 3' non-translated region

**SEQ ID NO: 8**

**[0048]**

Name: I389/NS3-3'/5.1

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-7767: HCV polyprotein from non-structural protein 3 to non-structural protein 5B of the cell culture-adapted mutant 5.1

**THIS PAGE BLANK (USPTO)**

5. 1813-3708: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
6. 3709-3870: Non-structural protein 4A (NS4A); NS3 protease cofactor
7. 3871-4653: Non-structural protein 4B (NS4B)
8. 4654-5994: Non-structural protein 5B [*sic*] (NS5A)
9. 5995-7767: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
- 7771-8001: HCV 3' non-translated region

**SEQ ID NO: 9**

**[0049]**

Name: I389/core-3'/5.1

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-10842: HCV polyprotein from core to non-structural protein 5B of the cell culture-adapted mutant 5.1
5. 1813-2385: HCV core protein; structural protein
6. 2386-2961: Envelope protein 1; structural protein
7. 2962-4050: Envelope protein 2; structural protein
8. 4051-4239: Protein p7
9. 4240-4890: Non-structural protein 2 (NS2); HCV NS2-3 protease
10. 4891-6783: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
11. 6784-6945: Non-structural protein 4A (NS4A); NS3 protease cofactor
12. 6946-7728: Non-structural protein 4B (NS4B)
13. 7729-9069: Non-structural protein 5A (NS5A)
14. 9070-10842: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
15. 10846-11076: HCV 3' non-translated region

**SEQ ID NO: 10**

**[0050]**

Name: I389-NS3-3'/19

Structure (nucleotide positions):



**THIS PAGE BLANK (USPTO)**

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-7767: HCV polyprotein from non-structural protein 3 to non-structural protein 5B of the cell culture-adapted mutant 19

5. 1813-3708: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
6. 3709-3870: Non-structural protein 4A (NS4A); NS3 protease cofactor
7. 3871-4653: Non-structural protein 4B (NS4B)
8. 4654-5994: Non-structural protein 5A (NS5A)
9. 5995-7767: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
- 7771-8001: HCV 3' non-translated region

**SEQ ID NO: 11**

**[0051]**

Name: I389/core-3'/19

Structure (nucleotide positions):

1. 1-341: HCV 5' non-translated region
2. 342-1193: HCV core protein-neomycin phosphotransferase fusion protein; selectable marker
3. 1202-1812: Internal ribosome binding site of the encephalomyocarditis virus; permits translation of the HCV open reading frame beyond it
4. 1813-10842: HCV polyprotein from core to non-structural protein 5B of the cell culture-adapted mutant 19
5. 1813-2385: HCV core protein; structural protein
6. 2386-2961: Envelope protein 1; structural protein
7. 2962-4050: Envelope protein 2; structural protein
8. 4051-4239: Protein p7
9. 4240-4890: Non-structural protein 2 (NS2); HCV NS2-3 protease
10. 4891-6783: Non-structural protein 3 (NS3); HCV NS3 protease/helicase
11. 6784-6945: Non-structural protein 4A (NS4A); NS3 protease cofactor
12. 6946-7728: Non-structural protein 4B (NS4B)
13. 7729-9069: Non-structural protein 5A (NS5A)
14. 9070-10842: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
15. 10846-11076: HCV 3' non-translated region

**THIS PAGE BLANK (USPTO)**

[0052] The invention will be explained in detail below with reference to examples of embodiment and associated tables and figures. The figures mentioned show:

**Fig. 1 A:** The structure of an HCV-RNA construct pursuant to the invention. At the very top is a schematic representation of the structure of the complete parental HCV genome with the positions of the genes for the cleavage products core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B within the polyprotein, and the 5' and 3' non-translated regions (5' NTR and 3' NTR) shown as horizontal bars, and with the two positions selected for producing the subgenome constructs, namely the position of the 'GDD catalytic domain' of NS5B RNA polymerase (GDD) and the position of the 3' boundary of the HCV-IRES (nucleotide positions 1 to 377 or 1 to 389) drawn above the genome diagram. The numbers below the genome diagram designate the corresponding nucleotide positions.

Below this are shown schematic representations of the structures of two modified HCV RNA constructs (subgenomes) pursuant to the invention, consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (Neo<sup>R</sup>), the EMC-IRES (E-1), and the HCV sequences from NS2 or NS3 to the authentic 3' end. The position of the 10-amino-acid deletion comprising the NS5B polymerase GDD unit is labeled in each case with a triangle ( ).

**Fig. 1 B:** The result of denaturing formaldehyde-agar gel electrophoresis for detecting replicated plus-strand RNA in transfected sub-passed Huh-7 cell clones.

The positions of the HCV-specific RNAs (arrows) and of the 28S rRNA are indicated to the right of Track 12, and the sizes (numbers of nucleotides) of the RNA markers (M) are indicated to the left of Track 1.

**Fig. 1 C:** The result of a PCR test followed by Southern Blot to demonstrate the absence of integrated replicon DNA in most of the selected cell clones. Tracks 1 and 2 show the positive controls, Track 13 shows the negative control. The numerical data to the left of Track 1 designate the size of the nucleotide marker molecule.

**Fig. 2 A:** The result of a PCR test followed by Southern Blot for the sensitive exclusion of integrated replicon DNA (plasmid molecule I<sub>377</sub>/NS3-3'/wt) in a cell clone (9-13) containing HCV RNA construct. Tracks 7 to 11 represent the result of a titration of DNA molecules of the construct I<sub>377</sub>/NS3-3'/wt without addition of total DNA of the cell clone 9-13, and Tracks 2-6 represent the same plasmid molecule with the addition of 1 g 9-13 DNA in each case prior to PCR (to exclude a PCR inhibitor in the DNA preparation). Track 13 represents the negative control (PCR without DNA probe). Track 1 shows the result obtained with one g total DNA from the cell clone 9-13.

**Fig. 2 B:** The result of a Northern Blot test to quantify HCV plus- and minus-strand RNA. The arrows label the positions of replicon RNA. The "plus" and "minus" data designate the positive (plus) and negative (minus) polarity of the RNA controls that were

**THIS PAGE BLANK (USPTO)**

applied to the gel. "Minus-strand" and "plus-strand" designate the specificity of the radioactive RNA probes.

**Fig. 2 C:** Result of formaldehyde-agar gel electrophoresis after radioactive labeling of the intracellularly replicated HCV RNA to detect resistance of the HCV RNA replication to dactinomycin.

**Fig. 3 A:** Detection of HCV-specific antigens in the selected cell clones by immunoprecipitation after metabolic radioactive labeling. Tracks 7-9 represent authentic size markers (that Huh-7 cells contained after transient expression of an HCV RNA construct); identified HCV PROTEINS ARE LABELED AT THE LEFT EDGE OF Track 1, and the molecular weights (in kilodaltons) are indicated at the right edge of Track 9.

**Fig. 3 B:** Results of an immunofluorescence test to determine the subcellular location of HCV antigens.

**Fig. 4:** Schematic representation of the structure of a selectable HCV RNA construct (complete genome) pursuant to the invention consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (NeoR), a heterologous IRES element, e.g. of the encephalomyocarditis virus (E-I), the complete HCV reading frame, and the authentic 3' NTR.

**Fig. 5:** Schematic representation of the structure of HCV RNA constructs with inserted antibiotic resistance gene (A) within the nucleotide sequence that codes for the polyprotein (monocistronic RNA), and (B) within the 3' NTR (bicistronic RNA).

**Fig. 6:** Schematic representation of the structure of HCV RNA constructs with inserted reporter gene (A) as part of an HCV replicon from NS3 to NS5B; the reporter protein is ultimately cleaved from the polyprotein by viral or cellular proteases and the selectable marker gene (selection gene) or the resistance gene is inserted in the cells by cotransfection, (B) as part of a fusion gene of a resistance gene and a reporter gene (for example for neomycin phosphotransferase and green fluorescent protein), (C) as part of a replicon of a resistance gene and a reporter gene (for example for neomycin phosphotransferase and green fluorescent protein), that are connected through a nucleotide sequence that codes for an amino acid sequence (hatched region), which can be cleaved by a protease or that has self-cleaving (autocatalytic) activity, (D) as an independent gene (green fluorescent protein in this case) that is expressed by its own internal ribosome-binding site (IRES); the resistance gene (in this case the neomycin phosphotransferase gene) is likewise expressed independently of this by its own internal ribosome-binding site (IRES) (polycistronic construct).

**Fig. 7:** Schematic representation of the structure of an HCV RNA construct in which the resistance gene is joined to the HCV RNA sequence through a ribozyme or a recognition site for a ribozyme. The heavy lines represent the HCV 5' and 3' NTRs, E-I is a heterologous internal ribosome-binding site that is necessary for the expression of the

**THIS PAGE BLANK** (USPTO

resistance gene, and the gray square represents the ribozyme or a recognition site for a ribozyme.

**Fig. 8:** Schematic representation of the structure of an HCV RNA construct with resistance gene and integrated foreign gene.

**Fig. 9:** Procedural method for comparing the specific infectiousness (expressed as the number of cell colonies formed) of total RNA versus in vitro transcripts. HCV RNA is prepared by in vitro transcription of a corresponding RNA construct and is quantified by measuring the optical density at 260 nm (OD 260 nm). A given number of these molecules is mixed with a definite amount of total RNA from natural Huh-7 cells and this mixture is inserted into natural Huh-7 cells using electroporation. In parallel to this, the total RNA of a cell clone that was prepared by the method described in Figure 1 is isolated by a method known in the state of the art and the amount of HCV RNA that it contains is determined by Northern Blot using an HCV-specific RNA probe followed by quantification by phospho-imager. A given amount of this total RNA is transfected into natural Huh-7 cells similarly to the in vitro transcripts. These cells in the two formulations are then subjected to G418 selection and the number of colonies formed is determined by counting after fixing and staining with Coomassie Brilliant Blue. To determine the transfection efficiency, 1  $\mu$ g of a plasmid that permits the expression of luciferase is added to each transfection formulation. An aliquot of the transfected cells is harvested after 24 hours and the luciferase activity in the particular cell lysate is determined. The number of colonies in each case is normalized to the luciferase expression.

**Fig. 10:** Sequence analysis of 9-13 clones. Total RNA of the cell clone 9-13 that was formed by transfection of the HCV RNA construct I377/NS3-3' was isolated by a method known in the state of the art, and the HCV RNA construct was amplified from nucleotide position 59 to 9386 by 'long-distance RT-PCR' using the primer S59 and A9413. The PCR fragments were cloned and 11 clones (named 9-13 A-K) were completely sequenced, with the clones D and I, E and G, and H and J proving to be identical. The positions of the amino acid differences in the NS3-5B region between the recloned HCV RNA and the parental construct are labeled with a heavy vertical line at the particular clone. Each clone was digested with the restriction enzyme *Sfi* and the particular fragment was inserted into the parental construct. Each of these clones was transfected in Huh-7 cells and the cells were subjected to selection as described in Figure 1. The number of cell clones obtained with each construct is noted on the right next to the particular construct.

**Fig. 11 A:** Principle of determination of replication using a reporter gene. The upper part of the figure illustrates the HCV DNA construct I<sub>389</sub>/Luc/NS3-3', consisting of the HCV 5' NTR (nucleotide positions 1-389), the luciferase gene (*luc*), the IRES of the encephalomyocarditis virus, HCV NS3-5B, and the 3' NTR. The position of the active center of NS5B RNA polymerase into which an inactivated amino acid replacement has been inserted is indicated as 'GND'. The plasmids that code for the replication-competent

**THIS PAGE BLANK** (USPTO)

or the defective HCV RNA construct are digested with the restriction enzyme *Sca* I and used for in vitro transcription with T7 RNA polymerase. After removal of the matrix DNA, the particular HCV RNA constructs are inserted into natural Huh-7 cells and these are harvested at regular intervals.

**Fig. 11 B:** Comparison of luciferase activities in cells transfected with the parental HCV RNA construct  $I_{389}/\text{Luc}/\text{NS3-3'}/\text{wt}$  (wt) or the following variants: Inactive RNA (318 DN), variant 9-13F, or variant 5.1. The cells were harvested 6 (not shown), 24, 48, 72, 96, 120, 144, and 168 hours after transfection and the luciferase activity was determined luminometrically.

**Fig. 12:** Selectable full-length HCV genomes (constructs  $I_{389}/\text{core-3'}/5.1$  and  $I_{389}/\text{core-3'}/9-13\text{F}$ ).

(A) Schematic representation of the full-length construct. The region between the two indicated recognition sites for the restriction enzyme *Sfi* I corresponds to the sequences of the highly adapted RNA variants 5.1. or 9-13F.

(B) Number of colonies that were obtained after transfection into HUH-7 cells of 0.1 g RNA transcribed in vitro of the constructs  $I_{389}/\text{core-3'}/5.1$  shown under A. The result of a representative experiment is shown.

(C) Detection of autonomously replicating full-length HCV RNAs in G418-resistant cell clones that were obtained after transfection of the corresponding in vitro transcript. The depiction shows the autoradiogram of a Northern Blot that was hybridized with a probe against the *neo*-resistance gene and the HCV 5' NTR. The controls shown in Tracks 1 and 2 each correspond to  $10^8$  molecules of the indicated in vitro transcripts, mixed with total RNA from natural Huh-7 cells. The negative control contains only total RNA from natural Huh-7 cells (Track 3). Tracks 4-9 contain 3-10 g total RNA from G418-resistant cell clones that were obtained after transfection of in vitro transcribed  $I_{389}/\text{core-3'}/5.1$  RNA and  $I_{389}/\text{core-3'}/9-13\text{F}$  RNA, respectively. The G418 concentration used for selection is indicated in each case. Five of the illustrated cell clones contain the highly adapted RNA variant 5.1 (Tracks 4-8), and one contains the adapted RNA variant 9-13F (Track 9).

**Fig. 13:** HCV RNA constructs with a reporter gene. (A) Bicistronic HCV RNA constructs. The reporter gene is translated using a separate IRES. (B) Monocistronic HCV RNA constructs. The reporter gene product is expressed as a fusion protein with an HCV protein. The two portions are joined through a recognition sequence for a viral or cellular protease that permits proteolytic separation of the two fused protein sections. In the example shown, the reporter gene product and the particular HCV protein are fused through a recognition sequence for ubiquitin (Ub).

**Fig. 14:** Tricistronic full-length HCV RNA construct that has a foreign gene inserted in addition to the resistance gene.

**THIS PAGE BLANK** (USPTO,

**Fig. 15:** Monocistronic HCV RNA constructs in which the resistance gene product is expressed as a fusion protein with the HCV portion. The resistance gene (RG) is either active as a fusion protein or it is fused with a proteolytically cleavable sequence in such a way that the resistance gene product is cleaved off from the HCV component by a cellular or viral protease. In the example shown, the resistance gene was fused with the particular HCV component through the sequence coding for ubiquitin (Ub).

### **Example 1: Preparation of HCV RNA constructs**

#### **(A) Synthesis and cloning of a complete HCV consensus genome by RT-PCR**

[0053] The HCV genome was isolated from the liver of a chronically infected patient; i.e. the HCV was isolated as described below:

[0054] The complete RNA was isolated from about 100 mg of liver according to the method of Chomczynski and Sacci (1987, Anal. Biochem. 162, 156). A reverse transcription was performed with 1 µg of this isolated RNA with the primer A6103 (GCTATCAGCCGGTTCATCCACTGC) or A9413 (CAGGATGGCCTATTGG CCTGGAG) and the 'expand reverse transcriptase' system (Boehringer Mannheim, Germany) according to the manufacturer's instructions. A polymerase chain reaction (PCR) was performed with the products of this reverse transcription (RT), using the 'expand long template' system (Boehringer Mannheim, Germany), with the buffer containing 2% dimethyl sulfoxide being used. After one hour at 42°C, 1/8 of this reaction mixture was used in a first PCR pass with the primers A6103 and S59 (TGTCTTCACGCAGAAAGCGTCTAG) or A9413 and S4542 (GATGAGCT CGCCGCGAAGCTGTCC) being used. After 40 cycles, 1/10 of this reaction mixture was used in a second PCR pass with the primers S58 and A4919 (AGCACAGCCCGCGTCATAGCACTCG) or S4542 and A9386 (TTAGCTCCCCG TTCATCGGTTGG). After 30 cycles the PCR products were purified by preparative agar gel electrophoresis and the eluted fragments were ligated into the vector pCR2.1 (Invitrogen) or pBSK II (Stratagene). Four clones of each fragment were analyzed and sequenced, and a consensus sequence was determined. For this purpose, the DNA sequences were compared with one another. The positions at which the sequence of one of the fragments differed from the others, were considered to be unwanted mutation. In the case of ambiguities of the sequence, shorter overlapping PCR fragments of the region in question were amplified and several clones were sequenced. In this way numerous potential mutations in each fragment could be identified and thus an isolate-specific consensus sequence could be established. This established consensus sequence or this genome belongs to the genotype 1b of worldwide distribution. The non-translated region at the 3' end (= 3' NTR) was obtained by conventional PCR, with an antisense primer being used that covers the last 24 nucleotides of the 'X-tail' known from the state of the art (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; and Rice, PCT/US 96/14033). The authentic non-translated region at the 5' end (= 5' NTR) downstream from the T7 promoter was produced by PCR, with an oligonucleotide being used on the one hand that corresponds to a shortened T7 promoter (TAA TAC GAC TCA CTA TAG) and

**THIS PAGE BLANK (USPTO)**

the first 88 nucleotides of HCV, and on the other hand one of the aforementioned plasmids that carries one of the 5' fragments of the genome. From the subgenomic fragments with the smallest number of non-consensus replacements, a complete HCV consensus genome was assembled and inserted into a modified pBR322 vector. Differences from the consensus sequence were eliminated by "site-directed mutagenesis". To prepare "run-off" transcripts with an authentic 3' end, the 3' NTR of the isolate (with the end TGT) was modified to AGT (according to the sequence of Genotype 3 = Clone 'WS' according to Kolykhalov et al., 1996, J. Virol 70, 3363) and an additional nucleotide replacement was also performed at position 9562 to retain the A:T base pairing in the hairpin structure at the 3' end of the 3' NTR (Kolykhalov et al. *ibid.*). To eliminate an internal restriction site for the enzyme *Scal*, a so-called silent nucleotide replacement was also performed. After the combination of the full-length genome with fitting 5' and 3' NTRs, the complete HCV sequence was checked. No unwanted nucleotide replacement was found.

[0055] The HCV genome prepared in this way should be hepatotropic by definition.

#### (B) Synthesis of selectable HCV subgenome constructs

[0056] Using the consensus genome described under (A), HCV subgenome constructs were prepared that contain the antibiotic resistance gene neomycin-phosphotransferase (NPT) and two sequences of internal ribosome binding sites (IRES). The biochemical procedural techniques needed to do this are known and familiar to one skilled in the art (see Sambrook, J., E. F. Fritsch, T. Maniatis, 1989, *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.; Ausubel et al. (eds.); 1994, *Current Protocols in Molecular Biology*, Vols. 1-3, John Wiley & Sons Inc., New York). The antibiotic resistance gene was inserted directly beyond the 5' NTR, so that a bicistronic RNA was obtained (see Fig. 1 A). However, the antibiotic resistance gene can also just as well be inserted at other sites on the HCV subgenome construct, for example in the nucleotide sequence coding for the polyprotein, whereby a monocistronic RNA is obtained (see Fig. 5 A) or in the 3' NTR (see Fig. 5 B). The IRES elements are one of the two HCV IRES variants nucleotides 1-377 or nucleotides 1-389, on the one hand, and on the other the IRES of the encephalomyocarditis virus that controls the translation of the HCV sequence downstream from the genes for NS2 or NS3 to the authentic 3' end of the genome.

[0057] The two HCV IRES variants mentioned were determined as follows: Based on deletion analyses of the 3' boundary of the HCV IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010), various sections of the 5' NTR were fused with the NPT gene and analyzed for the maximum number of colonies formed with reference to cotransfections with a plasmid containing the T7 RNA polymerase gene. The best results were obtained with the HCV sequences of 1-377 and 1-389. Since the AUG start codon of the HCV polyprotein is located at position 342, and is thus contained in the IRES sequence, there is fusion of 12 or 16 amino acids of the HCV capsid protein (core protein) with neomycin phosphotransferase (see Fig. 1 A).



**THIS PAGE BLANK (USPTO)**

[0058] These modified HCV subgenome constructs accordingly were given the designations I<sub>377</sub>/NS2-3' (or I<sub>377</sub>/NS3-3') and I<sub>389</sub>/NS2-3' (or I<sub>389</sub>/NS3-3'). They are illustrated schematically in Fig. 1A.

[0059] Various cell lines and primary cell cultures of human hepatocytes were transfected with in vitro transcripts of these modified parental HCV subgenome constructs I<sub>377</sub>/NS2-3' (or I<sub>377</sub>/NS3-3') and I<sub>389</sub>/NS2-3' (or I<sub>389</sub>/NS3-3').

[0060] As parallel negative controls for all transfection experiments, a correspondingly modified but defective subgenome was constructed for each modified parental HCV subgenome construct, which differs from the parental one in that it has a deletion of 10 amino acids that comprises the active center of NS5B RNA polymerase (Behrens et al., 1996 *EMBO J.* 15, 12; and Lohmann et al., 1997, *J. Virol.* 71, 8416).

#### (C) Synthesis of selectable HCV genome constructs

[0061] An NS2-3' subgenome construct that is joined at the 5' end to a fragment of the luciferase gene and the complete EMCV IRES was restricted with NcoI and SpeI, and purified by preparative agar gel electrophoresis. The vector thus obtained was ligated in a 3-factor ligation with an NcoI/NotI HCV fragment corresponding to nucleotide positions 342 to 1968 of the HCV genome, and with an NotI/SpeI fragment corresponding to nucleotide positions 1968-9605. The construct obtained, in which the complete HCV reading frame and the 3' NTR are downstream from the luciferase fragment and the EMCV IRES, was then restricted with PmeI and SpeI and ligated with the analogously restricted I<sub>389</sub>/NS3-3'/wt subgenome construct vector. This selectable HCV genome construct is illustrated in Fig. 4.

#### (D) Preparation of in vitro transcripts corresponding to the HCV RNA constructs

[0062] The purified plasmid DNAs described above were linearized with Scal and after phenol/chloroform extraction and isopropanol precipitation they were used in an in vitro transcription reaction using the following components: 80 mM HEPES, pH 7.5, 12.5 mM MgCl<sub>2</sub>, 2 mM Spermidin, 40 mM dithiothreitol, 2 mM of each NTP, 1 unit of RNasin/μl, 50 μg/ml restricted DNA, and about 2 units/μl of T7 RNA polymerase. After 2 hours at 37°C, half of the amount of T7 polymerase was added and the reaction formulation was incubated 2 hours longer. To remove DNA, the mixture was extracted with acidic phenol (U. Kedzierski, J. C. Porte, 1991, *Bio Techniques* 10, 210), precipitated with isopropanol, the pellet was dissolved in water, and was incubated with DNase (2 units per μg DNA) for 60 min. at 37°C. After the subsequent extraction with acidic phenol, acidic phenol/chloroform, and chloroform, and precipitation with isopropanol, the dissolved RNA was quantified by optical density measurement and the absence of damage to it was checked by formaldehyde-agar gel electrophoresis.

#### Example 2: Transfection experiments with the hepatoma cell line Huh-7

**THIS PAGE BLANK (USPTO)**

[0063] Care was taken in all transfection experiments to see that any matrix DNA had previously been removed, to avoid such DNA being integrated into transfected cells and being able to impart neomycin resistance to them independently of HCV replication. For this reason, after the in vitro transcription (Example 1 D), the reaction mixture was treated with 2 units of DNase per  $\mu\text{g}$  of DNA for 60 min. at  $37^\circ\text{C}$  and extracted with acidic phenol, acidic phenol/chloroform, and chloroform. Before use for transfection, the precipitated RNA was analyzed by formaldehyde-agar gel electrophoresis.

[0064] Three separate transfection experiments were carried out with the highly differentiated human hepatoma cell line Huh-7 (by the method of Nakabayashi et al. 1982, *Cancer Res.* 42, 3858). In each case,  $15\ \mu\text{g}$  of RNA was inserted into  $8 \times 10^6$  Huh-7 cells using electroporation, and these cells were then plated in culture dishes 10 cm in diameter. 24 hours later neomycin (= G418) was added at a final concentration of 1 mg/ml. The culture medium was changed twice each week. After 3-5 weeks, small colonies were recognizable and were isolated and subjected to another pass under the same culture conditions.

[0065] The cell clones that were obtained in the course of the first experiment were isolated and submitted to subpassage. During this procedure, most of the clones died, and the final yield was only 9 clones of cells that had been transfected with a defective HCV genome construct, namely a defective NS2-3' HCV RNA. Other than a shortened doubling time and the occasional appearance of irregularly formed cells, no constant morphological differences were found between these 9 cell clones and the one cell clone (clone 8-1) or the parental Huh-7 cells.

[0066] The main criteria for the functioning HCV genome constructs are the formation of viral RNA of correct size and the absence of (integrated) plasmid DNA, which could transfer or impart G418 resistance.

[0067] To determine the HCV RNA in the Huh-7 cells, the total RNA was isolated and analyzed by the current Northern Blot method using a plus strand-specific riboprobe (= RNA probe). To do this, total RNA from the particular cell clones were isolated by the method of Chomczynski and Sacchi 1987, *Anal. Biochem.* 162, 156, and  $10\ \mu\text{g}$  of RNA, which corresponds to the total RNA content of  $0.5-1 \times 10^6$  cells, was separated out by denaturing formaldehyde-agar gel electrophoresis (Tracks 3 to 12 of Fig. 1 B). At the same time, as size markers with authentic sequence,  $10^9$  in vitro transcripts (ivtr.) that correspond to the  $I_{389}/\text{NS2-3'}/\text{wt}$  or  $I_{389}/\text{NS3-3'}/\text{wt}$  replicon RNAs were separated along with it (Track 1 and Track 2, respectively). The separated RNA was transferred to Nylon membranes and hybridized with radioactively labeled plus strand-specific RNA probe that was complementary to the complete NPT gene and the HCV IRES from nucleotide 377 to nucleotide 1. The positions of the HCV-specific RNAs (arrows) and of the 28S rRNA are indicated to the right of Track 12, and the sizes (numbers of nucleotides) of the RNA markers are indicated to the left of Track 1. The RNA marker fragments contain HCV sequences and therefore hybridize with the riboprobe (= RNA probe). The results of this analysis are shown in Fig. 1 B.

**THIS PAGE BLANK (USPTO)**

[0068] With the exception of Clone 8-1 transfected with the defective clone 8-1, all of the cell clones provided homogeneous HCV RNAs of correct length (about 8640 nucleotides in the case of NS2-3' and about 7970 nucleotides in the case of the NS3-3' replicon). This finding indicates that the functional replicons and the functional HCV genome constructs transmit G418 resistance. To exclude the possibility that the G418 resistance can be attributed to a plasmid DNA that is integrated into the genome of the Huh-7 host cell and is transcribed under the control of a cellular promoter, the DNA from each clone was examined by NPT-gene-specific PCR. The DNA from the selected Huh-7 cell lines was isolated by digestion with proteinase K (40 µg/ml, 1 h, 37°C) in 10 mM Tris, pH 7.5, 1 mM EDTA, 0.5% SDS, and then extraction with phenol, phenol/chloroform, and precipitation with isopropanol. The DNA precipitate was dissolved in 10 mM Tris (pH 7.5) and 1 mM EDTA and incubated for 1 hour with Rnase A. After phenol/chloroform extraction and ethanol precipitation, 1 µg DNA, corresponding to  $4-8 \times 10^4$  cells, was analyzed by PCR using NPT-gene-specific primer (5'-TCAAGACCGACCTGTCCGGTGCCC-3' and 5'-CTTGAGCCTGGCGAACAGTTCGGC-3'), and a DNA fragment consisting of 379 nucleotides was produced. The specificity of the PCR product was tested by the Southern Blot method, with a digoxigenin-labeled DNA fragment being used that corresponds to the NPT gene. As positive controls (to detect any contaminating nucleic acids present), the PCR procedure was carried out with  $10^7$  plasmid molecules or 1 µg DNA from a BHK cell line that was stably transfected with a neomycin resistance gene, and as negative control the PCR was carried out with the same reagents but with no added DNA.

The results of this study are shown in Fig. 1 C. Tracks 1 and 2 represent the positive controls, and Track 13 represents the negative control. The numerical data to the left of Track 1 designate the size of the nucleotide marker molecule. Other than in clone 7-3 (Figure 1C, Track 3), which originates from cells after transfection with an NS2-3' replicon/NS2-3' HCV genome construct, and in clone 8-1 (Fig. 1C, Track 12), which originates from cells after transfection with a defective HCV genome, NPT DNA was not detectable in any of the cell clones. This finding is further indication that the G418 resistance of most of the clones was imparted by the replicating HCV RNA. However, independently of these results, it is also improbable that HCV RNAs with the correct size are produced from integrated plasmid DNA, since the plasmids used for the in vitro transcription contained neither a eucaryotic promoter nor a polyadenylation signal. In the case of the clone 7-3, the resistance was therefore with the greatest probability imparted both by the HCV RNA construct and the replicating HCV RNA, and also by an integrated NTP DNA sequence, while the resistance of the cells of clone 8-1 can be attributed exclusively to the integrated plasmid DNA.

[0069] To verify that the G418 resistance is imparted by an autonomously replicating HCV RNA, clone 9-13 (Fig. 1 B, Track 11) was subjected to further tests. Clone 8-1, which carries integrated copies of the NPT gene, was used everywhere as the negative control. With the objective of rigorously excluding the presence of NPT DNA in clone 9-13, a PCR was performed that permits the detection of < 1000 NPT gene copies in ~40,000 cells. The result of this PCR is shown in Fig. 2A. In detail, the procedure for this PCR was as follows:

**THIS PAGE BLANK (USPTO)**

In each case  $10^6$ - $10^2$  plasmid molecules (I<sub>377</sub>/NS3-3'/wt) were used either directly (Tracks 7-11) in the test or after adding 1 µg of 9-13 DNA (Tracks 2-6). The specificity of the amplified DNA fragment was determined by Southern Blot using an NPT-specific probe. A PCR with no DNA probe was performed as a negative control (Track 12).

Even with this sensitive method, no plasmid DNA was found in one µg of DNA of cell clone 9-13 (Track 1). To estimate the amount of HCV plus- and minus-strand RNAs in these cells, a dilution series of total RNA was analyzed by the Northern Blot method using a plus- or minus-strand-specific radioactively labeled riboprobe (= RNA probe). To do this, 8, 4, or 2 µg of total RNA that had been isolated from the cell clones 9-13 and 8-1 was analyzed in parallel with known amounts of analogous in vitro transcripts with plus- or minus-strand polarity (control RNAs) by the Northern Blot method, and then subjected to hybridization. The hybridization was carried out with a plus-strand-specific riboprobe that covered the complete NPT gene and the HCV IRES ('plus strand', top image table) or with a minus-strand-specific RNA probe that was complementary to the NS3 sequence ('minus strand', bottom image table). The arrows mark the positions of replicon RNA. The results of this analysis are shown in Fig. 2 B.

In the case of the plus strand, about  $10^8$  copies of total RNA were detected, which corresponds to 1000-5000 HCV RNA molecules per cell, while the amount of minus strand RNA was lower by a factor of 5 to 10. This result is in agreement with the assumption that the minus strand RNA is the replicative intermediate form or intermediate copy that serves as a template for the synthesis of the plus strand molecule.

Since the reaction is catalyzed essentially by the viral RNA-dependent RNA polymerase, the synthesis of HCV RNAs should be resistant to dactinomycin, an antibiotic that selectively inhibits RNA synthesis from DNA matrices but not RNA synthesis from RNA matrices. To confirm this supposition, cells were incubated with [<sup>3</sup>H]uridine in the presence of dactinomycin, the radioactively labeled RNAs were extracted, separated by means of denaturing agar gel electrophoresis, and analyzed with a commercial bio-imager using an [<sup>3</sup>H]-sensitive image plate. To do this, in each case about  $5 \times 10^5$  cells of clones 9-13 and 8-1 were incubated with 100 µCi of [<sup>3</sup>H]uridine for 16 hours in the absence (-) or presence (+) of 4 µg/ml dactinomycin (Dact). After this labeling reaction, the total RNA was prepared and analyzed by formaldehyde-agar electrophoresis. In the two first tracks, only 1/10 of the total RNA is shown. The radioactively labeled RNA was made visible with a BAS-2500 Bio-imager (Fuji Company).

The results of this analysis are shown in Fig. 2 C. In conformity with the inhibitor profile of NS5B polymerase (Behrens et al., 1996, *EMBO J.* 15, 12, AND Lohmann et al., 1997, *J. Virol.* 71, 8416), the replication of the HCV RNA had not been affected by dactinomycin, while the synthesis of cellular RNA had been inhibited. To verify the identity of the viral RNA, an RT PCR was performed to reclone the replicated sequences. Sequence analysis of the recloned RNA showed that the RNA in clone 9-13 is HCV-specific and conforms to the transfected transcript of the HCV construct I<sub>377</sub>/NS3-3'/wt.

**THIS PAGE BLANK (USPTO)**

[0070] To analyze the viral proteins, the cells in question were first radioactively labeled metabolically with [ $^{35}\text{S}$ ]methionine/cystein, then lysed, and then the HCV-specific proteins were isolated from the cell lysates by immunoprecipitation. The results of these analyses are shown in Fig. 3 A. In detail, the procedures was as follows: Cells of cell clones 9-13 (wt) and 8-1 ( ) had been radioactively labeled metabolically by treatment for 16 hours with a protein-labeling mixture (e.g. NEN Life Science) familiar to one skilled in the art and commercially available. The HCV-specific proteins had been separated out of the cell lysate by immunoprecipitation (IP) under non-denaturing conditions (for example according to Bartenschlager et al., 1995, *J. Virol.* 69, 7519) and using three different antisera (3/4, 5A, 5B as marked at the top ends of Tracks 1 to 12). The immunocomplexes were analyzed by means of Tricine SDS-PAGE and visualized by autoradiography. To obtain authentic size markers, the homologous replicon construct I<sub>377</sub>/NS3-3'/wt was subjected to transient expression with the Vaccinia Virus T7 hybrid system in Huh-7 cells. The products obtained had been treated as size markers (Tracks 7-9) in parallel with the cells of clones 9-13 and 8-1. Identified HCV proteins are labeled at the left edge of Track 1, and the molecular weights (in kilodaltons) are indicated at the right edge of Track 9. It should be noted that the NS3/4-specific antiserum ('3/4') used reacts preferentially with NS4A and NS4B, which leads to under-representation of NS3.

[0071] All of the viral antigens were clearly detectable and their apparent molecular weights showed no difference from those that were determined after transient expression of the same bicistronic HCV RNA constructs in the original Huh-7 cells. To determine the subcellular distribution of the viral antigens, an immunofluorescence detection reaction was carried out using NS3- and NS5A-specific antisera (for example by the method of Bartenschlager et al., 1995 *J. Virol.* 69, 7519). To do this, cells of clones 9-13 (wt) and 8-1 ( $\Delta$ ) were fixed with methanol/acetone for 24 hours after plating on cover glasses, and were incubated with polyclonal NS3- or NS5A- specific antisera. The bound antibodies were visualized with a commercially available FITC-conjugated anti-rabbit antiserum. To suppress nonspecific fluorescence signals, the cells were counter-stained with the dye 'Evans Blue'.

[0072] The results of this detection test are shown in Fig. 3 B. Strong fluorescence was detected in the cytoplasm with both antisera. The NS5A-specific antisera also led to weak cell nucleus fluorescence, which indicates that at least small amounts of this antigen also reach the cell nucleus. The generally dominating presence of the viral antigens in the cytoplasm, however, is a strong indication that the HCV RNA replication takes place in the cytoplasm, just as is the case with most RNA viruses.

[0073] These results demonstrate clearly that the experimental formulation described here succeeded in constructing a cell culture system for HCV whose efficiency exceeds everything known heretofore by orders of magnitude, and for the first time permits the detection of viral nucleic acids and proteins by conventional and proven biochemical methods. Only this efficiency permits any detailed studies at all of HCV pathogenesis, genetic analyses of various HCV functions, and a precise study of virus/host cell interactions, by which new points of attack can be defined for the development of antiviral therapy.

**THIS PAGE BLANK (USPTO)**

**Example 3: Transfection of Huh-7 cells with HCV genome constructs**

[0074] Huh-7 cells are transfected and selected as described in Example 2, but in this case selectable constructs are used that contain the complete virus genome. The cell clones obtained are tested analogously to Example 2 by PCR for the absence of HCV DNA, and the productive replication of HCV RNA is then detected by Northern Blot, [<sup>3</sup>H]uridine labeling in the presence of dactinomycin, and detection of viral proteins and antigens, preferably using Western Blot, immunoprecipitation, or immunofluorescence. In contrast to the formulations described in Example 2, the construct described here can also contain complete and very probably infectious viruses, which is not the case in the subgenome constructs described there (in Example 2). These viruses that are present in the cell and the cell culture supernatant are concentrated, for example by ultracentrifugation, immunoprecipitation, or precipitation with polyethylene glycol, and all exogenous nucleic acids, i.e. those not incorporated in the virus particle, are digested by incubation with nucleases (RNase, DNase, micrococcus nuclease). In this way all contaminating nucleic acids that are not contained in the protecting virus particle can be removed. The protected viral RNA is isolated after inactivation of the nucleases, for example by incubation with proteinase K in a buffer containing SDS, by extraction with phenol and phenol/chloroform, and detected by Northern Blot or RT-PCR using HCV-specific primer. The combination of the described HCV consensus genome with a selection marker is also critical in the experimental formulation for the efficient production of viral RNA, viral protein, and with it HCV particles.

**Example 4: Preparation and use of an HCV RNA construct in which the resistance gene is joined to the HCV subgenome sequence through a ribozyme or a recognition site for a ribozyme**

[0075] An HCV RNA construct is prepared according to Example 1 or Example 3, in which an antibiotic resistance gene is joined to the HCV RNA sequence through a ribozyme or a recognition site for a ribozyme. Such constructs are illustrated schematically in Fig. 7. Huh-7 cells are transfected with this HCV RNA construct as described in Example 2. After transfection into the cells, selection with the corresponding antibiotic is first carried out. The cloned-in ribozyme in the cell clones thus obtained is activated, or in the case of a construct that carries a recognition site for a ribozyme, the ribozyme is inserted into the cell (for example by transfection of a ribozyme construct or infection with a viral expression vector in which the corresponding ribozyme was used). In both cases the resistance gene is separated from the HCV RNA sequence by ribozyme-mediated cleavage. The result in the case of the HCV genome construct is an authentic HCV genome with no resistance gene, which is capable of forming authentic infectious virus particles. In the case of the HCV subgenome construct, an HCV replicon with no resistance gene is formed.

**Example 5: Cotransfection of an HCV RNA construct with a separate luciferase transfection construct**

**THIS PAGE BLANK (USPTO**

[0076] An HCV RNA construct is prepared according to Example 1 (A) or Example 3 or Example 4. In parallel with this, a transfection construct is prepared that comprises the luciferase gene, with this luciferase gene being joined by means of a first nucleotide sequence that codes for an HCV protease cleavage site (e.g. NS3 protease cleavage site), to a second nucleotide sequence that codes for another protein or part of another protein. The HCV RNA construct and the transfection construct are inserted into any host cells, preferably hepatoma cells, especially Huh-7 cells. This can be done by the method described in Example 2. The product of the modified luciferase gene is a luciferase fusion gene in which the luciferase is inactive because of its fusion with the foreign part. In transfected cells with high HCV replication, the fusion protein, that does contain a cutting site for an HCV protease, is cleaved and thus the active form of luciferase is liberated, which can be determined by luminometric measurement. If the replication of the HCV RNA construct is inhibited, the fusion protein is not cleaved and no active luciferase is liberated. Consequently, the quantitative determination of luciferase is a measure of the replication of the HCV subgenome construct. A cellular protein that is inactivated or activated by the HCV proteins or nucleic acid can also be used as a so-called surrogate marker. In this case the expression or activity of this surrogate marker is a measure of the replication of the viral DNA.

**Example 6: Preparation of HCV subgenome constructs with integrated foreign gene for use as liver-specific gene ferries for gene therapy**

[0077] These recombinant and selectable HCV subgenome constructs are transfected into trans-complementing helper cell lines, i.e. into cell lines that express the missing functions (for example the structural proteins) inducibly or structurally. Cell clones that contain a functional HCV subgenome construct can be established by appropriate selection. The virus structural proteins expressed by the host cells permit the formation of virus particles into which the RNA of the HCV subgenome construct is inserted. The result is therefore virus-like particles that contain an HCV subgenome construct including the cloned-in foreign gene and that can transfer this to other cells by infection. An example of such a construct is shown in Fig. 8. There is also the ability to use the HCV subgenome construct pursuant to the invention with integrated foreign gene described here, directly as an expression vector. In this case, the procedure is analogous to the method described above, but with the difference that cell lines that express no trans-complementing factors are transfected. In this case the HCV construct therefore serves only as an expression vector.

**Example 7: Preparation of cell culture-adapted HCV RNA constructs**

**(A) Isolation procedure**

[0078] The procedure for determining adaptive mutations and for preparing cell culture-adapted HCV RNA constructs was as follows: Cells were transfected with an HCV RNA construct as described under Examples 1 and 2, and G418-resistant cell clones were prepared. To determine the replication competence (which means in this context the number of G418-resistant cell clones that are contained per microgram of transfected

**THIS PAGE BLANK (USPTO)**

HCV RNA or HCV RNA construct), by way of example, the total RNA was isolated from one of the cell clones called 9-13 (Fig. 1B, Track 11), and the amount of HCV RNA contained in it was determined by Northern Blot as described in Fig. 2 B. 10 micrograms of the total RNA, which contained about  $10^9$  molecules of HCV RNA, was then inserted into natural Huh-7 cells by electroporation (Fig. 9). In parallel with this,  $10^9$  in vitro transcripts of the analogous *neo*-HCV RNA that had been made up to a total RNA amount of 10  $\mu$ g with isolated total RNA from natural Huh-7 cells was transfected into natural Huh-7 cells. After selection with G418, the number of cell colonies, expressed as 'colony-forming units (cfu) per microgram of RNA' in the two formulations was determined. With a concentration of 500  $\mu$ g/ml of G418 in the selection medium, the number of colonies that were obtained with the HCV RNA contained in the *isolated* total RNA, amounted to about 100,000 cfu per microgram of HCV RNA. On the other hand, only 30-50 colonies were obtained with the same amount of in vitro transcribed HCV RNA. This result proves that the specific infectiousness of the HCV RNA that was isolated from the cell clones is about 1,000-10,000 as high as the infectiousness of the analogous in vitro transcripts. The procedural method is shown in Fig. 9.

[0079] Using 'long-distance RT-PCR', the HCV RNA from the total RNA of the 9-13 cells was amplified, the PCR amplificate was cloned, and numerous clones were sequenced. Comparison of the sequences of these recloned RNAs with the sequence of RNA that was originally inserted into the natural Huh-7 cells showed that the recloned RNAs had numerous amino acid replacements that were distributed over the entire HCV sequence (Fig. 10). *Sfi*I fragments of these recloned mutants were inserted into the original replicon construct to replace its analogous *Sfi*I fragment, and RNAs of the particular mutants were inserted into natural Huh-7 cells. After selection with G418, the number of colonies formed was then determined for each HCV RNA mutant. While only 30-50 colonies per microgram of RNA were obtained with the initial RNA, the number of colonies from two of the recloned variants was distinctly higher (Fig. 10). In the case of the HCV RNA constructs 9-13I and 9-13C, the specific infectiousness was 100-1,000 cfu per microgram of RNA, and for the 9-13F replicon it was even 1,000-10,000 cfu per microgram of RNA. These results show that the amino acid replacements in the analyzed NS3-5B region of mutants 9-13I, 9-13C, and especially 9-13F led to a clear increase of replication competence. On the other hand, none of the other HCV RNA constructs (9-13 A, B, G, H, or K) were still replication competent, and therefore contained lethal mutations.

[0080] To answer the question of which of the amino acid replacements in the 9-13F construct led to the increase in replication, the replacements were inserted individually or in combination into the initial HCV RNA construct, and the corresponding RNAs were inserted into natural Huh-7 cells. The result of transfection with these RNAs is summarized in Table 1. It is apparent from this that the high replication competence in the present example is caused by multiple mutations. The amino acid replacements in the HCV RNA sections NS5A and NS4B make the largest contribution. The individual replacements in the NS3 region also make a contribution that may originate from synergism of these individual replacements.

**THIS PAGE BLANK (USPTO)**

These findings show that there was enrichment of HCV RNAs that had clearly higher replication competence from the G418 selection of the cells that were transfected with the *neo*-HCV RNA constructs. HCV RNA constructs with very diverse replication efficiencies can be selected with the experimental formulation described here. The higher the concentration of antibiotic in the selection medium in/on which the cells containing the HCV RNA construct are cultured for purposes of selection, the higher the degree of adaptive mutations has to be, and thus the higher the replication efficiency in the HCV RNA constructs in question has to be, for the cells to be able to grow. If the selections are carried out with low concentrations of antibiotic, cells can also survive and multiply that have fewer adaptive mutations and lower replication efficiency.

The HCV RNA construct 9-13F described above, which contained multiple adaptive mutations, had demonstrated higher replication efficiency than the parental HCV RNA. To obtain HCV RNAs with still higher replication in cell culture, the HCV RNA that was contained in the total RNA of a chosen cell clone was passed repeatedly into Huh-7 cells. This chosen cell clone, called 5-15, was obtained by transfection with the RNA construct I<sub>389</sub>/NS3-3' (Fig. 1). It corresponds largely to cell clone 9-13, which was prepared by transfection with an HCV RNA construct that had an HCV IRES shorter by 22 nucleotides (I<sub>377</sub>/NS3-3'; Fig. 1). 10 micrograms of total RNA isolated from the cell clone 5-15 was inserted into natural Huh-7 cells by electroporation, and the cells were subjected to selection with 1 mg/ml of G418. From one of the cell clones produced in this way, total RNA was in turn isolated, transfected into natural Huh-7 cells, and selected analogously. This procedure was repeated a total of four times. After the fourth passage, the total RNA was isolated from one cell clone and the *neo*-HCV RNA was amplified using 'long-distance RT-PCR'. The amplified DNA fragment was digested with the restriction enzyme SfiI and inserted into the SfiI-restricted initial construct I<sub>389</sub>/NS3-3'. More than 100 DNA clones were obtained in all, and were first analyzed by restriction digestion. In vitro transcribed RNA from each of about 80 of these clones was inserted into natural Huh-7 and subjected to selection with 500 mg/ml of G418. Of the 80 *neo*-HCV RNAs studied, almost all of them proved to be replication-defective. With two mutants, however, called 5.1 and 19, their specific infectiousness, expressed as 'colony-forming units' per milligram of RNA, was very clearly increased (Table 2). By multiple passage of the RNA in cell culture, HCV RNAs can apparently be prepared whose replication efficiency because of mutations (so-called "adaptive mutations") is higher by several orders of magnitude than that of the RNA originally cloned from the patient.

### (B) Modification procedure

[0081] Such adaptive mutations produced and identified according to (A) can be transferred to an HCV RNA construct with low replication competence and lead to a massive increase of replication of this construct. This increase is so great that it can be shown with them that HCV RNAs can be caused to replicate in cell culture that no longer have any selectable marker gene. Fig. 12 shows a comparison of the replication efficiency of HCV RNAs that correspond either to the initial sequence or to the adapted sequences 9-13F or 5.1. For the purpose of easier measurement, the *neo*-gene was removed and replaced by the gene for luciferase. An HCV RNA construct that was



**THIS PAGE BLANK (USPTO)**

replication-defective because of an inactivating mutation of the NS5B RNA polymerase served as negative control. Even 24 hours after transfection, a distinct difference was found in luciferase activity between the defective RNA and the 9-13F or 5.1 constructs, while hardly any difference was found between the defective RNA (318 DN) and the initial RNA construct (wt) that had no adaptive mutations. The highest luciferase activity and with it the highest replication during the entire period of observation was obtained with the 5.1 RNA. These findings not only demonstrate the high replication efficiency of this RNA, but they also show that it is possible with adapted HCV RNA constructs to construct a cell culture system for which the presence of a selectable gene is no longer necessary. Table 3 gives a summary review of the nucleotide and amino acid differences between the initial construct and the mutants 9-13F, 5.1, and 19.

### **Example 8: Preparation of cell culture-adapted full-length HCV RNA genomes**

[0082] In Examples 1 to 7 a subgenomic HCV RNA was used that lacked the entire structural protein region from core up to and including p7 or NS2. In the present Example 8, it will be shown that it is possible with the help of the adapted NS3-5B sequence to cause a full-length HCV genome to replicate in cell culture. For this purpose, the Sfil fragment of the highly adapted HCV RNA 5.1 prepared according to Example 7 was first transferred into a selectable full-length HCV genome (Fig. 12). This HCV genome was transfected into natural Huh-7 cells and subjected to selection with various G418 concentrations. Depending on the strength of selection (the G418 concentration), different numbers of cell clones were obtained (Fig. 12 B). In comparison with this, no colonies were obtained with the unchanged full-length HCV genome that contained no adaptive mutations, as well as with the negative control that was replication-defective because of an inactivating mutation in the NS5B RNA polymerase. To demonstrate that the cell clones thus formed actually contained an autonomously replicating full-length HCV genome, total RNA from several clones was isolated and analyzed by Northern Blot. The full-length HCV was clearly detectable in all of the cell clones (Fig. 12). This proves unequivocally that it is possible by using the HCV sequences adapted to cell cultures to prepare a full-length HCV genome that replicates at high efficiency and autonomously in a cell line, in other words adapted full-length HCV genomes can also be prepared with the system pursuant to the invention. Since this clone furthermore has the complete HCV sequence, including the structural proteins necessary for the formation of virus particles, it is possible with this system to prepare large amounts of infectious virus particles in cell cultures. To detect these viruses, cell-free supernatants of cells that carry a replicating full-length HCV genome are added to natural Huh-7 cells, and the cells thus infected are subjected to selection with G418. Each cell clone that grows under these conditions is attributed to an infected cell. The viruses in the cell culture supernatants of cells that have a replicating full-length HCV genome, however, can also be concentrated and purified by various methods known in the state of the art such as ultracentrifugation or microdialysis, and can then be used to infect natural cells. This procedure shows clearly that cell culture-adapted full-length HCV genomes can be prepared with the HCV cell culture system pursuant to the invention that replicate in cells with high efficiency and that produce infectious viruses. These can likewise be detected by infection of an experimental animal, preferably chimpanzees.

**THIS PAGE BLANK (USPTO)**

**Example 9: Preparation of full-length HCV constructs and HCV subgenome constructs with reporter gene.**

[0083] An HCV RNA construct is prepared in which a reporter gene is inserted instead of the antibiotic resistance gene (Fig. 13). Replication in this case can be determined by the amount or activity of the reporter gene or of the reporter gene product. The reporter gene is preferably a gene from the group comprising the luciferase gene, the CAT gene (chloramphenicol-acetyl transferase gene), the lacZ gene (beta-galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene), or the SEAP gene (secreted alkaline phosphatase gene). These reporter genes and their products, namely the corresponding reporter proteins, can be determined by fluorescence, chemiluminescence, colorimetrically, or using immunological methods (e.g. enzyme-linked immunosorbent assay, ELISA). The reporter gene can be expressed either by its own IRES or in the form of a fusion protein that is either active as such or is joined to an HCV protein by means of a proteolytically cleavable amino acid sequence in such a way that it is split off from it by a cellular or viral (HCV) protease.

**Example 10: Preparation of full-length HCV constructs with integrated foreign genes for use as liver-specific gene ferries for gene therapy or as expression vectors.**

[0084] The construct (Fig. 14) is inserted into cells and leads there to the formation of HCV virus particles that can be used for the infection of other cells. Since the virus particles have encapsulated an RNA with a foreign gene, it can be used in the cells thus infected to produce the protein encoded by this foreign gene. Cells that have been transfected with the construct likewise express the foreign gene.

**Example 11: Preparation of monocistronic HCV RNA constructs in which the resistance gene product is expressed as a fusion protein with the HCV component**

[0085] For certain studies it is advantageous for the HCV RNA to have no heterologous IRES element. Examples of such studies are determinations of interferon resistance. If a cell that has an HCV RNA construct is incubated with alpha- or beta-interferon, there is a reduction of replication of the HCV RNA. To explain the mechanism of action it is necessary for the HCV RNA to have no heterologous IRES, since otherwise it cannot be determined whether the interferon-imparted inhibition is caused by inhibition of HCV replication or by inhibition of the heterologous IRES. For this reason, constructs are prepared in which the resistance gene is fused with an HCV protein (Fig. 15). Either the fusion protein is active as such, or the resistance gene product is joined to an HCV protein by means of a proteolytically cleavable amino acid sequence in such a way that it is split off from it by a cellular or viral (HCV) protease.

**THIS PAGE BLANK (USPTO)**

- 1) Specific infectiousness (cfu/ $\mu$ g RNA) of the HCV RNA constructs with adaptive mutations that were found in the 9-13F mutants and were inserted into the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt
- 2) Amino acid replacement<sup>1</sup>
- 3) None
- 4) all together

**THIS PAGE BLANK (USPTO)**

Table 2

Spezifische Infektiositäten (cfu/ $\mu$ g RNA) des parentalen HCV-RNA-Konstrukts I <sub>389</sub> /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte ante	RNA-Vari- ante	cfu/ $\mu$ g RNA <sup>1</sup>
3	Wildtyp	30 - 50
	9-13 C	100 - 1.000
	9-13 I	100 - 1.000
	9-13 F	1.000 - 10.000
	5.1	50.000 - 100.000
	19	50.000 - 100.000

<sup>1</sup> Colony-forming units (number of cell clones) in a selection from 500  $\mu$ g/ml G418.

- 1) Specific infectiousness (cfu/ $\mu$ g RNA) of the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and of the variants 9-13C, 9-13I, 9-13F, 5.1, and 19.
- 2) Transfected RNA variant
- 3) Wild type

**THIS PAGE BLANK (USPTO)**

**Table 3:** Nucleotide and amino acid differences between the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the mutants 9-13I, 9-13F, 5.1, and 19

HCV Mutante 1	Nukleotidposition 2	Nukleotidaustausch 3	Aminosäureaustausch h 4
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
	9346	T > C	Val > Ala
9-13 F	3866	C > T	-
	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
	8486	C > T	-
	8991	C > G	Arg > Gly
NK5.1	4180	C > T	Thr > Ile
	4679	C > T	-

1) HCV mutant

2) Nucleotide position

3) Nucleotide replacement

4) Amino acid  
replacement

**THIS PAGE BLANK (USPTO)**

	4682	T > C	-
	5610	C > A	Leu > Ile
	6437	A > G	-
	6666	A > G	Asn > Asp
	6842	C > T	-
	6926	C > T	-
	6930	T > C	Ser > Pro
	7320	C > T	Pro > Ser
	7389	A > G	Lys > Glu
NK19	3946	A > G	Glu > Gly
	4078	C > G	Ala > Gly
	4180	C > T	Thr > Ile
	4682	T > C	-
	5610	C > A	Leu > Ile
	5958	A > T	Met > Leu
	6170	T > A	-
	6596	G > A	-
	6598	C > G	Ala > Gly
	6833	C > T	-
	6842	C > T	-
	6930	T > C	Ser > Pro
	7141	A > G	Glu > Gly
	7320	C > T	Pro > Ser
	7389	A > G	Lys > Glu
	7735	G > A	Ser > Asn

[0086] The differences in nucleotide and amino acid sequences between the initial HCV RNA sequence Con 1 (EMBL gene bank No. AJ238799) and those of the cell culture-adapted HCV RNAs are indicated. The numbers refer to the nucleotide and amino acid positions of the HCV isolate Con 1.

**THIS PAGE BLANK (USPTO)**

**Patent Claims**

1. Hepatitis C virus (HCV) cell culture system that comprises essentially eucaryotic cells that contain inserted HCV-specific genetic material, characterized by the fact that

the eucaryotic cells are human hepatoma cells and that the inserted HCV-specific genetic material is an HCV RNA construct that comprises the HCV-specific RNA sections 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, and additionally a selectable marker gene (selection gene).

2. Cell culture system pursuant to Claim 1, characterized by the fact that

the hepatoma cells originate from a commercial hepatoma cell line.

3. Cell culture system pursuant to Claim 1, characterized by the fact that

the hepatoma cells are obtained from a hepatoma primary cell culture.

4. HCV RNA construct, characterized by the fact that

it comprises the HCV-specific RNA sections 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, and additionally a selectable marker gene (selection gene).

5. HCV RNA construct pursuant to Claim 4, characterized by the fact that

it comprises a nucleotide sequence according to the sequence protocols SEQ ID NO: 1 to SEQ ID NO: 11.

6. HCV RNA construct pursuant to Claim 4, characterized by the fact that

the 3' NTR has a nucleotide sequence that is selected from the group of nucleotide sequences (a) to (i) listed below:

(a) **ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT  
TTAGCTTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTG  
TTGCCCAGACTGGAGT**

1 043 399 A2

**THIS PAGE BLANK (USPTO)**

- (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTC TTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GGAGC
- (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTAATCTTT TTTTTTTCT TTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GCAGC
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTT TCTTTTTTT TGAGAGAGAG  
AGTCTCACTC TGTGCCCAG ACTGGAGT
- (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTT TCTTTTTTT TGAGAGAGAG  
AGTCTCACTC TGTGCCCAG ACTGGAGT
- (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTT TCTTTTTTT TGAGAGAGA  
GAGTCTCACT CTGTTGCCCA GACTGGAGT
- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTAGTCT TTTTTTTTT CTTTTTTTT GAGAGAGAGA  
GTCTCACTCT GTTGCCCAGA CTGGAGT
- (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTTAAT CTTTTTTTT TTTTCCTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT  
TTTTTTAATC TTTTTTTTT TTTTCTTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

7. HCV RNA construct pursuant to one of the claims 4 to 6, characterized by the fact that  
the selectable marker gene is a resistance gene and especially an antibiotic  
resistance gene.
8. HCV RNA construct pursuant to one of the claims 4 to 6, characterized by the fact that  
the selectable marker gene is a neomycin phosphotransferase gene.
9. HCV RNA construct pursuant to one of the claims 4 to 8, characterized by the fact that  
the selectable marker gene is integrated into the HCV RNA downstream from the  
5' NTR.

**THIS PAGE BLANK (USPTO)**

10. HCV RNA construct pursuant to one of the claims 4 to 9, characterized by the fact that

the selectable marker gene is joined to the HCV RNA through a ribozyme or a recognition site for a ribozyme.

11. HCV RNA construct pursuant to one of the claims 4 to 10, characterized by the fact that

it has an integrated reporter gene.

12. HCV RNA construct pursuant to claim 11, characterized by the fact that

the reporter gene is a gene from the group consisting of the luciferase gene, the CAT gene (chloramphenicol acetyltransferase gene), the lacZ gene (beta-galactosidase gene), the GFP gene (green fluorescence gene), the GUS gene (glucuronidase gene), and the SEAP gene (secreted alkaline phosphatase gene).

13. HCV RNA construct pursuant to one of the claims 4 to 11, characterized by the fact that

its replication affects the expression of a (cellular) surrogate marker gene.

14. HCV RNA construct pursuant to one of the claims 11 to 13, characterized by the fact that

the resistance gene is cloned into the open reading frame of the HCV RNA in such a way that it is convertible to an active form only after proteolytic processing.

15. HCV RNA construct pursuant to one of the claims 11 to 14, characterized by the fact that

the reporter gene and the selectable marker gene are arranged spatially in the construct in such a way that they jointly express a fusion protein.

16. Cell culture system pursuant to one of the claims 1 to 3, characterized by the fact that

the HCV RNA construct is a construct according to at least one of the claims 4 to 15.

17. Cell culture system pursuant to Claim 1, characterized by the fact that

the cells containing the HCV RNA construct are deposited with the DSMZ, Braunschweig, FRG, under the archive number DSM ACC 2394 (laboratory designation HuBl 9-13).

**THIS PAGE BLANK (USPTO**

18. Use of a cell culture system pursuant to one of the claims 1 to 3 or 16 to 17 and/or of an HCV RNA construct pursuant to one of the claims 4 to 15 for the preparation and/or evaluation and/or testing of therapeutic agents and/or diagnostic agents for the treatment in particular of HCV infections.

19. Use of a cell culture system pursuant to one of the claims 1 to 3 or 16 to 17 and/or of an HCV RNA construct pursuant to one of the claims 4 to 15 for the preparation of a vaccine against HCV infections.

20. Use of an HCV RNA construct pursuant to one of the claims 4 to 15 for the preparation of a liver cell-specific gene ferry for gene therapy.

21. HCV RNA construct pursuant to one of the claims 4 to 15, characterized by the fact that

it has an integrated foreign gene and is suitable for inserting this foreign gene into a target cell that is suitable for the expression of this foreign gene.

22. Method for obtaining cell culture-adapted mutants of an HCV RNA construct pursuant to one of the claims 4 to 15, with the mutants having higher replication efficiency than the HCV RNA construct, characterized by the fact that

a cell culture system pursuant to Claim 1 in which the inserted HCV-specific genetic material is an HCV RNA construct with selection gene according to one of the claims 4 to 15, is cultured on/in the selection medium corresponding to the selection gene, the grown cell clone is harvested, and the HCV RNA constructs or parts thereof are isolated from these cell clones.

23. Method pursuant to Claim 22, characterized by the fact that

the isolated HCV RNA constructs are again passed at least once, namely inserted into cells of a cell culture system pursuant to Claim 1, the cell culture system pursuant to Claim 1 in which the inserted HCV-specific genetic material is the isolated HCV RNA construct with selection gene is then cultured on/in the selection medium corresponding to the selection gene, the grown cell clone is harvested, and the HCV RNA constructs are isolated from these cell clones.

24. Method for preparing mutants of a full-length HCV genome or of a partial HCV genome or of any HCV construct with higher replication efficiency than the original full-length HCV genome or partial genome or HCV RNA construct, characterized by the fact that

a cell culture-adapted mutant of an HCV RNA construct is prepared and isolated by a method pursuant to Claim 22 or 23,

**THIS PAGE BLANK (USPTO)**

the nucleotide and amino acid sequences of these mutants are determined and the type, number, and positions of the nucleotide and amino acid mutations are determined by comparison with the nucleotide and amino acid sequence of the original HCV RNA construct, and these mutations are introduced into an (isolated) full-length HCV genome or a partial HCV genome or any HCV RNA construct either by selective mutagenesis or by replacement of sequence sections that contain the mutations in question.

25. Cell culture-adapted HCV RNA construct with high replication efficiency, characterized by the fact that

it is derivable by nucleotide and/or amino acid mutations of an HCV RNA construct pursuant to one of the claims 4 to 15 and it is obtainable by a method pursuant to one of the claims 22 to 24.

26. Cell culture-adapted HCV RNA construct pursuant to Claim 25, characterized by the fact that

it has one or more of the amino acid replacements listed below, namely 1283 arg -> gly and/or 1383 glu -> ala and/or 1577 lys -> arg and/or 1609 lys -> glu and/or 1936 pro -> ser and/or 2163 glu -> gly and/or 2330 lys -> glu and/or 2442 ile -> val.

27. Cell culture-adapted HCV RNA construct pursuant to Claim 25 or 26, characterized by the fact that

it has one or more of the nucleotide and/or amino acid replacements listed in Table 3, with Table 3 being part of this Claim.

28. Cell culture-adapted mutants of an HCV RNA construct or of a full-length HCV genome with higher replication efficiency than the original HCV RNA construct or than the original full-length HCV genome, characterized by the fact that

it is available by a method in which the type and number of mutations are determined in a cell culture-adapted HCV RNA construct pursuant to Claim 24 by sequence analysis and sequence comparison, and these mutations are introduced into an HCV RNA construct, particularly into an HCV RNA construct according to one of the claims 4 to 15, or into an (isolated) full-length HCV RNA genome, either by selective mutagenesis or by replacement of sequence sections that contain the mutations in question.

29. Hepatitis C virus particle or virus-like particle characterized by the fact that it is obtainable by a method pursuant to one of the claims 22-24.

30. Cells infected with hepatitis C virus particles or virus-like particles pursuant to Claim 29.

**THIS PAGE BLANK (USPTO)**

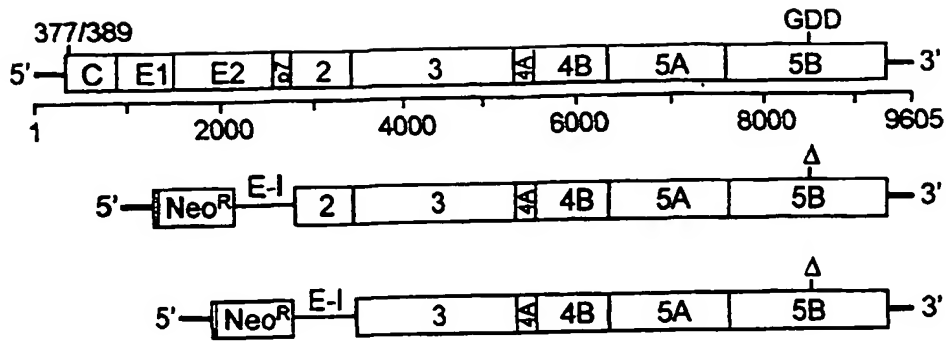
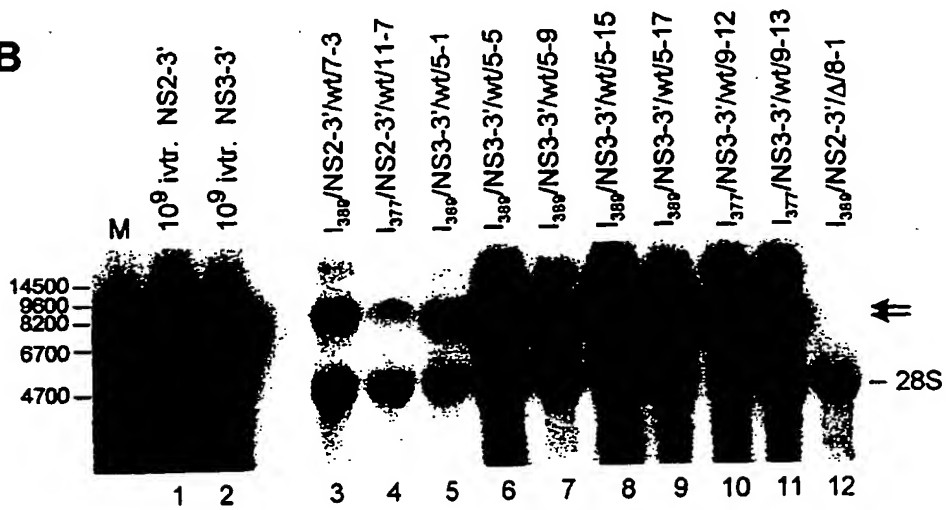
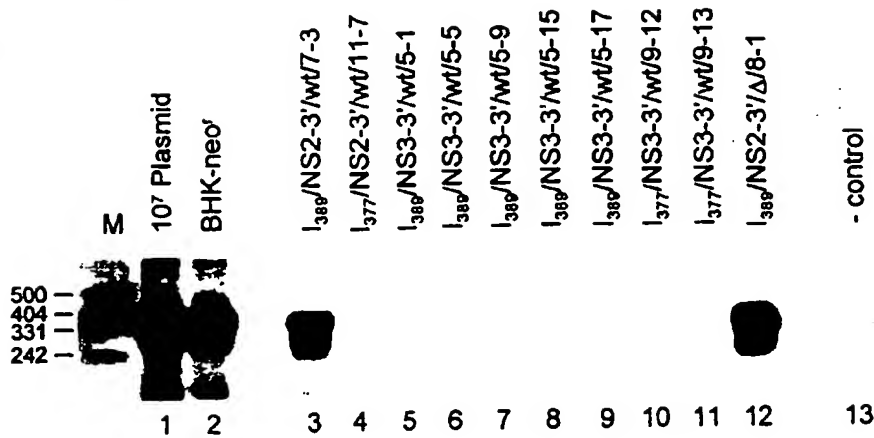
**A****B****C**

Fig. 1

**THIS PAGE BLANK (USPTO**

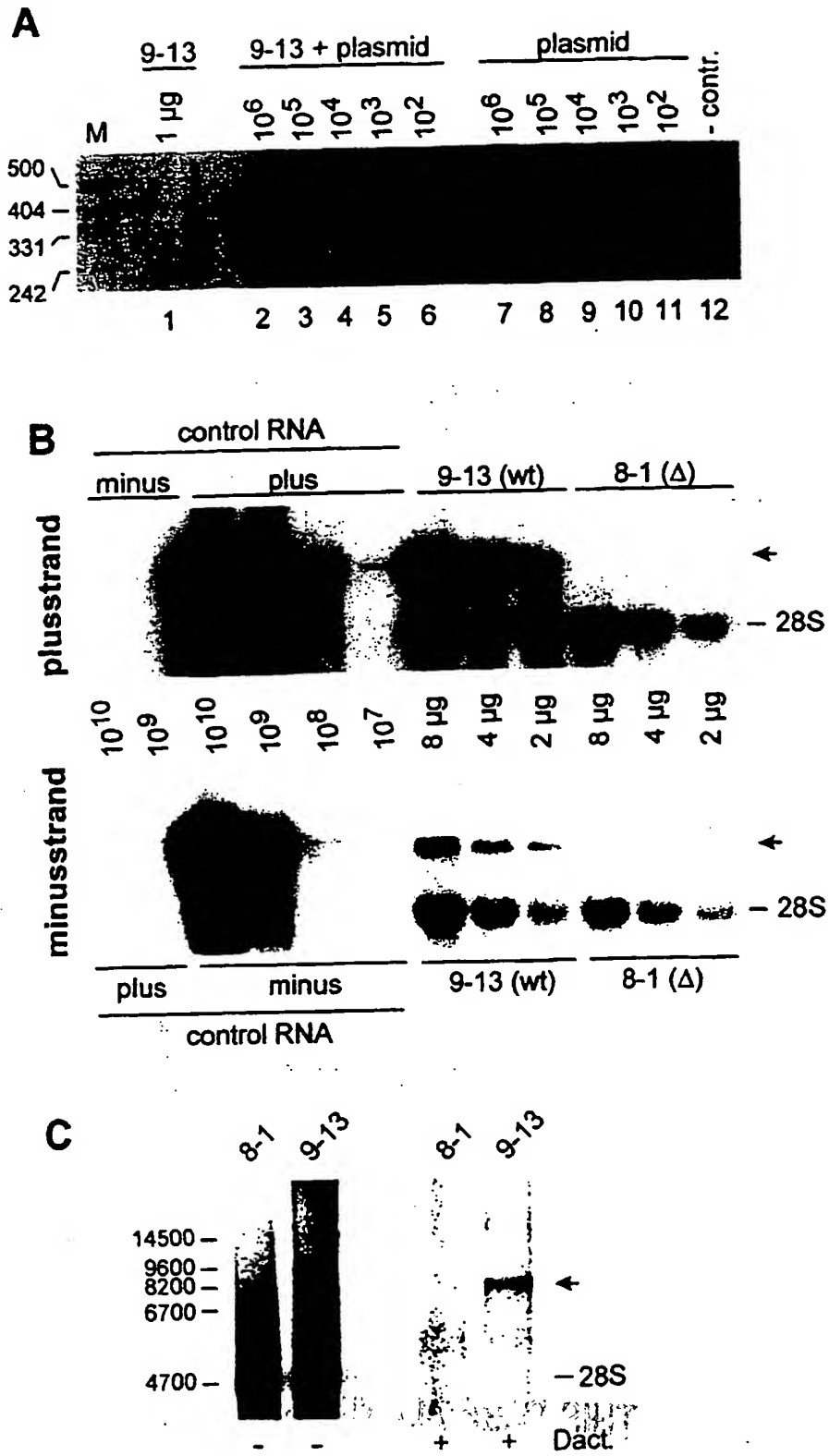


Fig. 2

**THIS PAGE BLANK (USPTO)**

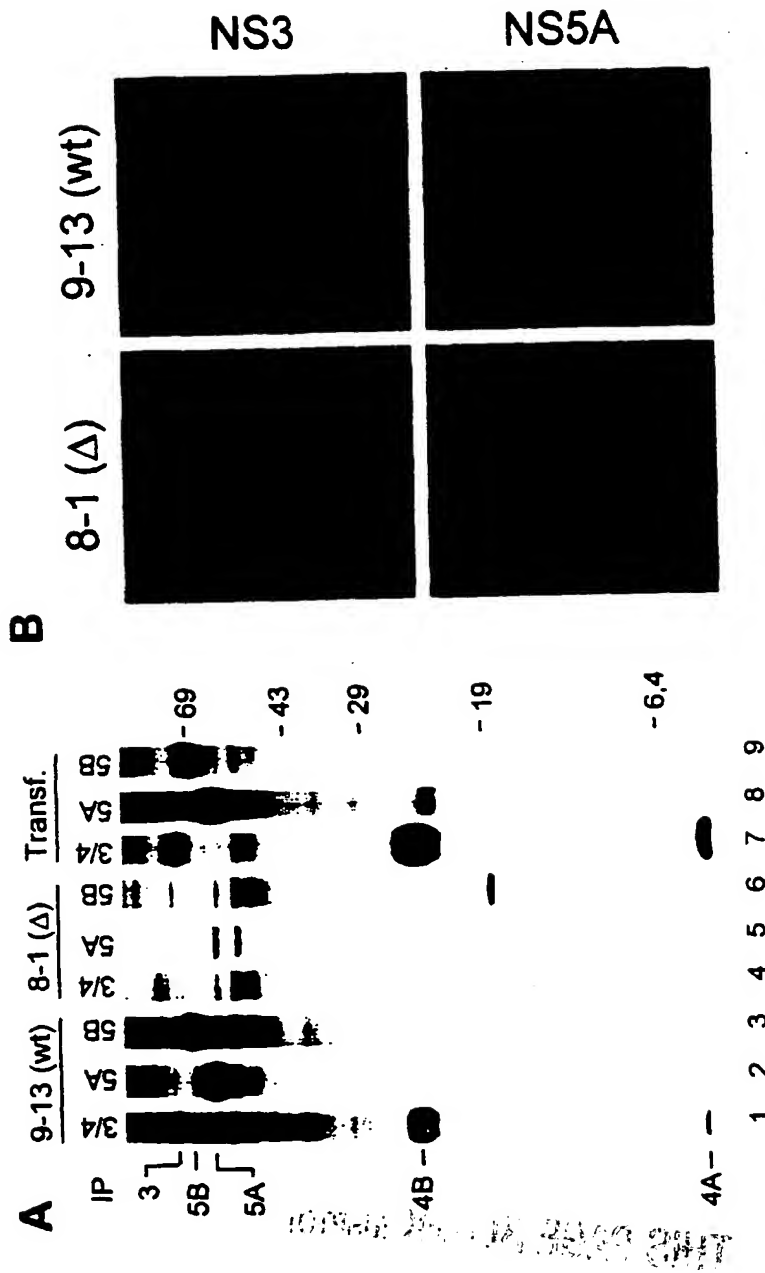


Fig. 3

**THIS PAGE BLANK (USPTO)**

[KEY TO FIGURES 5-8:]

FIGURE 5:

1) Resistance gene

FIGURE 6:

1) Reporter gene

FIGURE 7:

1) Ribozyme

2) (Recognition sequence)

FIGURE 8:

3) Therapeutic foreign gene

**THIS PAGE BLANK (USPTO)**

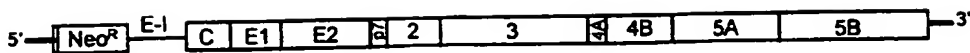


Fig. 4

A



B

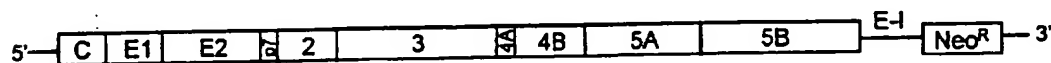
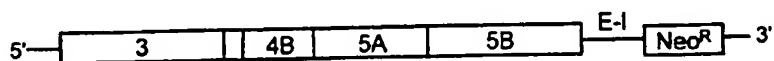


Fig. 5

10/10/2000 10:10:10

**THIS PAGE BLANK (USPTO)**

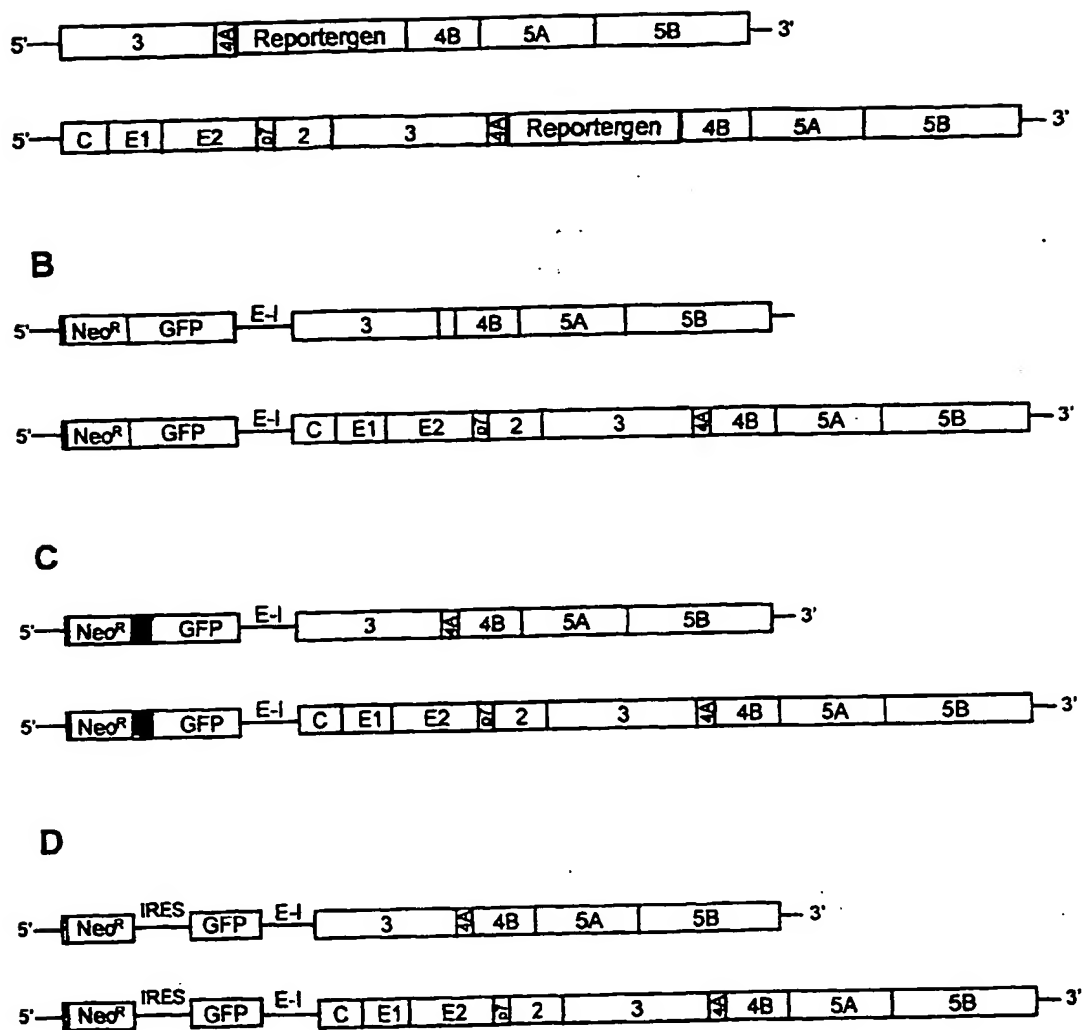


Fig. 6

**THIS PAGE BLANK (USPTO)**

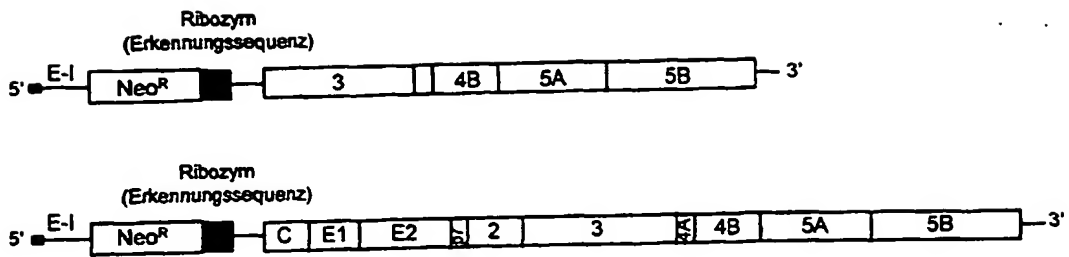


Fig. 7

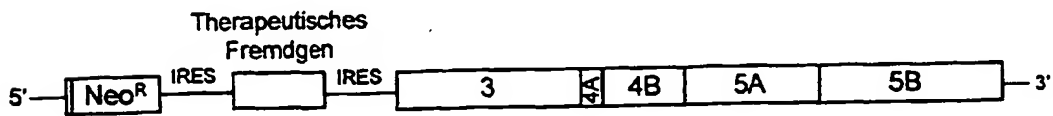


Fig. 8

**THIS PAGE BLANK (USPTO)**

[KEY TO FIGURE 9 ]:

- 1) In vitro transcript of cloned HCV RNA
- 2) G418-resistant colonies
- 3) Preparation of total cell RNA from individual colonies
- 4) Determination of number of molecules (OD 260 nm)
- 5) Total RNA from natural Huh-7 cells
- 6) Determination of the number of HCV RNA molecules (Northern Blot)
- 7) Total cell RNA with known number of HCV RNA molecules  
+ CMV luciferase plasmid (transfection control)**
- 8) Natural Huh-7 cells**
- 9) • Electroporation
- 10) • Plating out
- 11) • G418 selection
- 12) Cell colonies after about 3 weeks**

**THIS PAGE BLANK (USPTO**

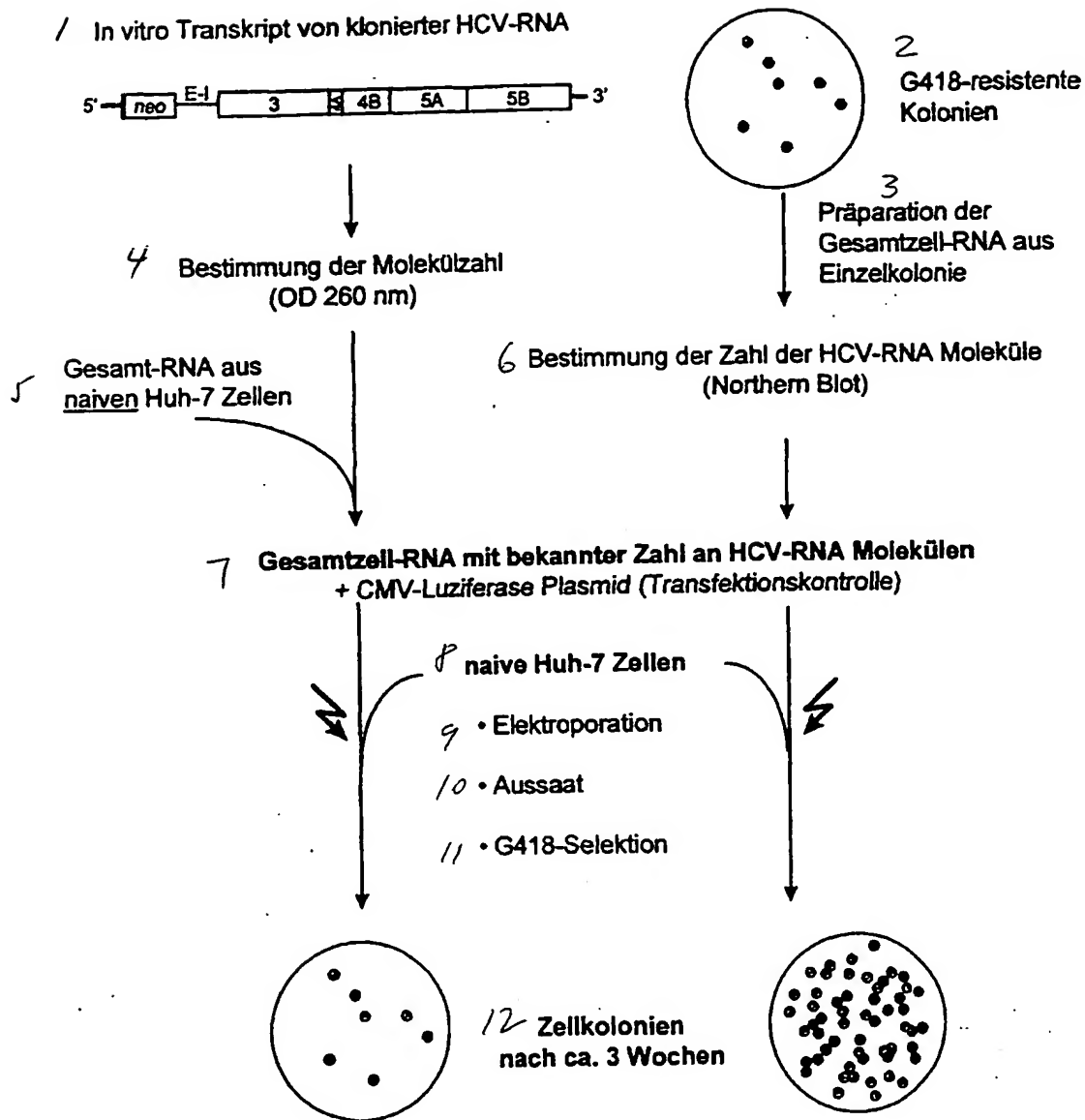


Fig. 9

**THIS PAGE BLANK (USPTO)**

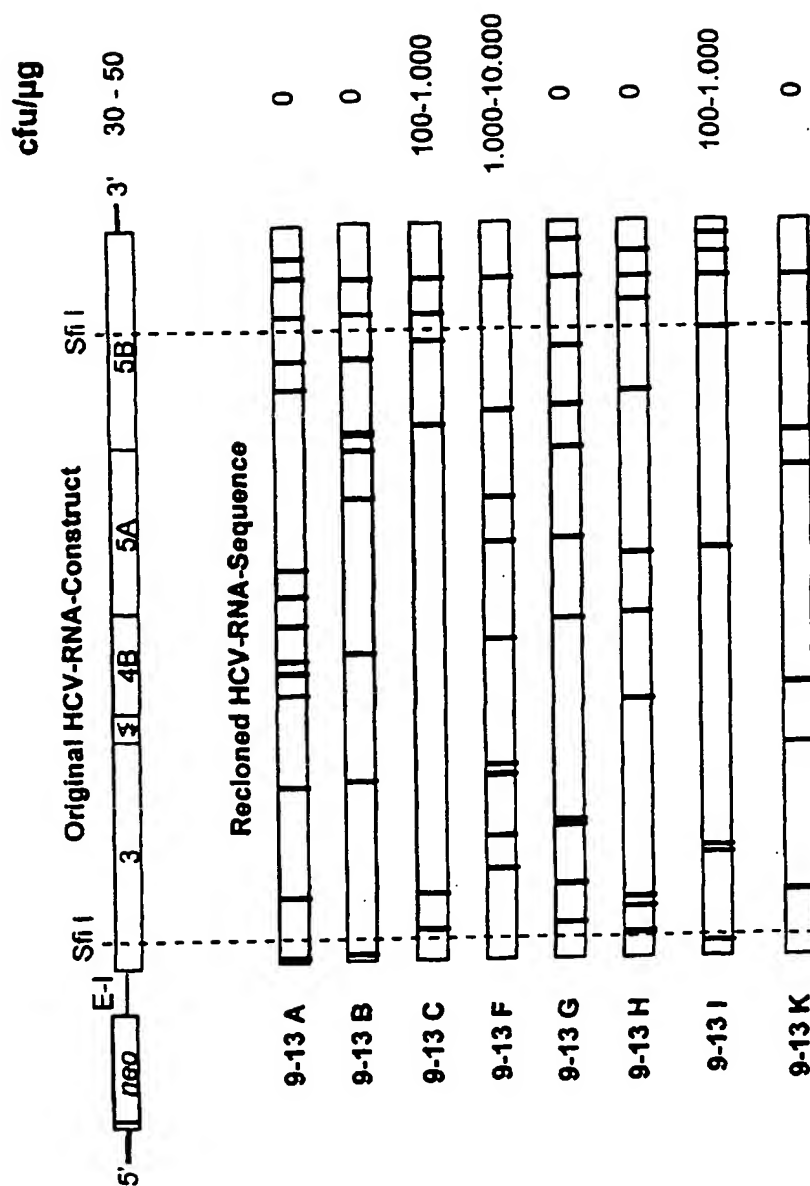


Fig. 10

**THIS PAGE BLANK (USPTO)**

[KEY TO FIGURE 11]:

- 1) • Linearization of plasmid
- 2) • in vitro transcription
- 3) • removal of matrix DNA
- 4) Wild type HCV RNA construct
- 5) Defective HCV RNA construct
- 6) • Electroporation into natural Huh-7
- 7) • Cell plating
- 8) • Cell harvests at regular intervals
- 9) • Determination of luciferase activity
- 10) Luciferase activity (RLU x  $10^4$ )
- 11) Time after transfection (hours)

**THIS PAGE BLANK (USE)**

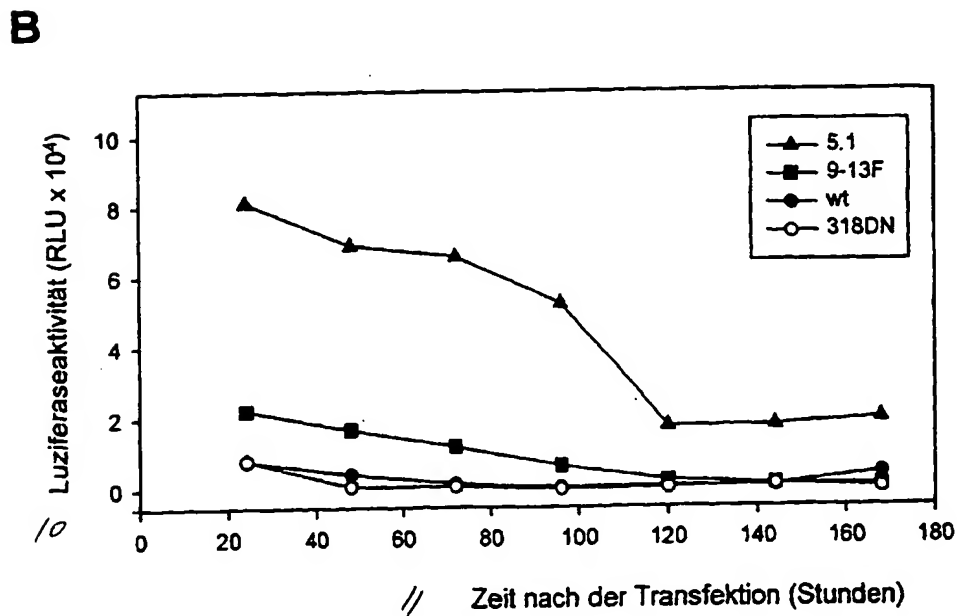
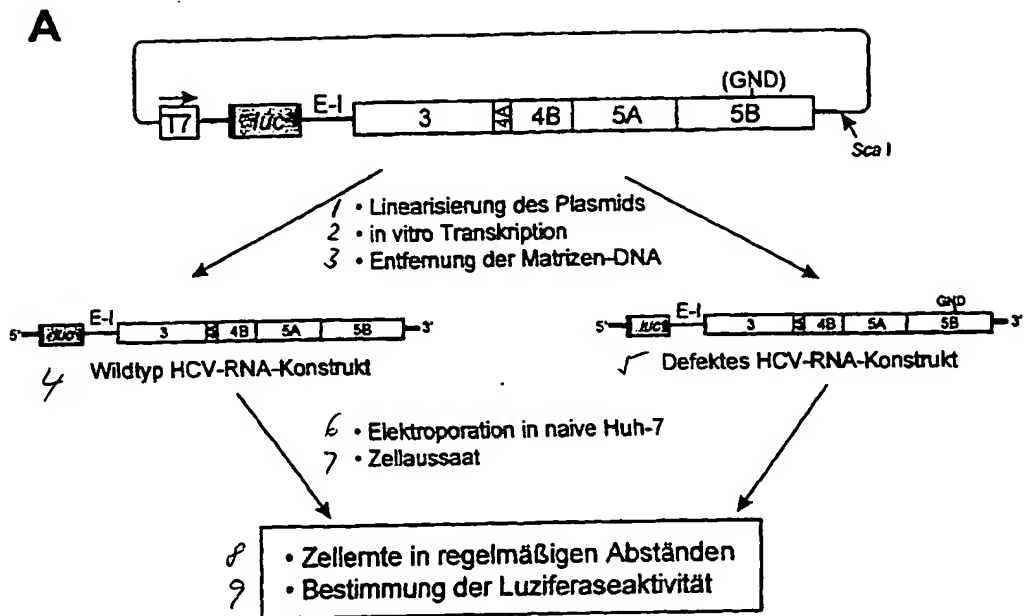


Fig. 11

**THIS PAGE BLANK (USPTO**

[KEY TO FIGURE 12]:

1) 5.1 or 9-13F mutants

2) G418 concentration

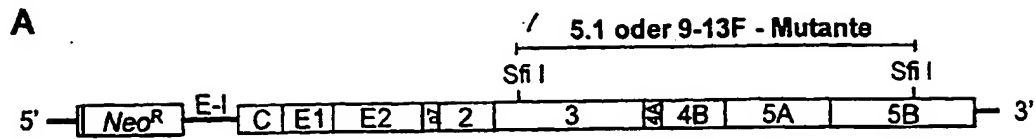
3) Number of colonies obtained:

4) Control

5) Negative control



**THIS PAGE BLANK (USPTO)**



**B**

2	G418-Konzentration:	500 µg/ml	250 µg/ml	100 µg/ml
3	Anzahl erhaltener Kolonien:	1	23	75

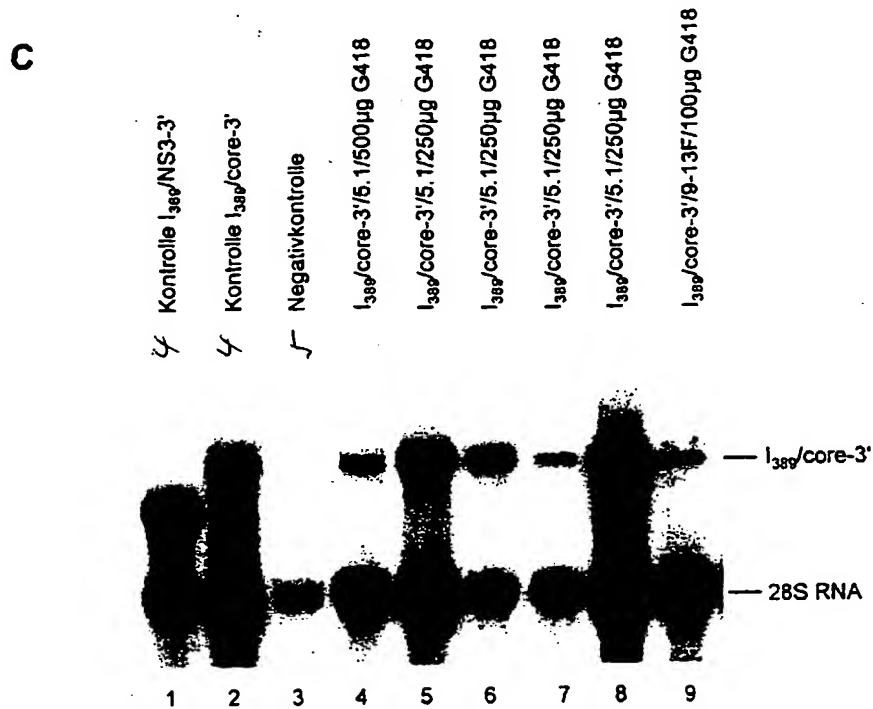


Fig. 12

**THIS PAGE BLANK (USPTO)**

[KEY TO FIGURES 13-14]:

FIGURE 13:

**1) Reporter genes**

FIGURE 14:

**2) Resistance genes**

ORIGINAL OF 3/11/2017

**THIS PAGE BLANK (USPTO)**

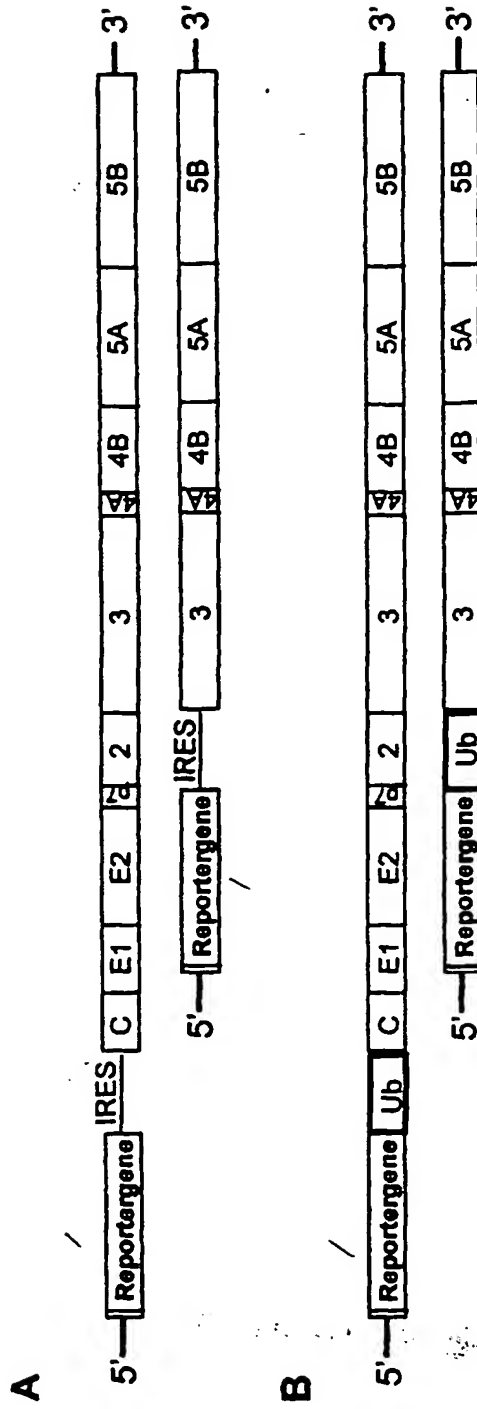


Fig. 13

**THIS PAGE BLANK (USPTO)**

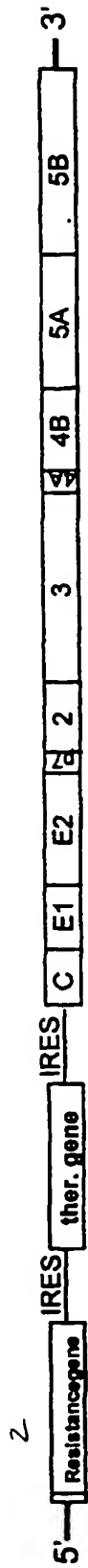


Fig. 14

**THIS PAGE BLANK (USPTO)**

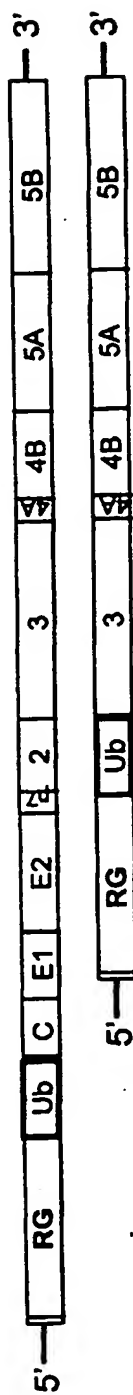


Fig. 15

**THIS PAGE BLANK** (USPTO

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**